

Innovation durch Konzeptübertragung. **Der Rückgriff auf Bekanntes bei der Erzeugung technischer Neuerungen** **am Beispiel der Multiagentensystem-Forschung**

Ingo Schulz-Schaeffer¹

TU Berlin, Institut für Soziologie, Sekr. FR 2-5, Franklinstr. 28/29, D-10587 Berlin

Zusammenfassung: Eine frühe soziologische Antwort auf die Frage danach, wie das Neue in die Welt kommt, lautet: im Wesentlichen durch Rekombination von Bekanntem. Eine auf wissenschaftliche Neuerungen bezogene Variante dieses Erklärungsmusters besagt, dass wissenschaftliche Innovationen in günstigen Fällen durch Übertragung bewährten Wissens aus einem gut beherrschten Forschungsgebiet auf ein anderes, bislang weniger gut beherrschtes Gebiet erfolgen. Grundlage hierfür ist eine angenommene Strukturgleichheit zwischen den Phänomenen der beiden Forschungsgebiete, die es erlaubt, etabliertes Wissen aus dem einen Forschungsgebiet in geeigneter modifizierter Form als theoretische Innovation in das andere Gebiet zu importieren. Der vorliegende Beitrag unternimmt den Versuch, dieses Konzept auf technische Innovationen zu übertragen. Es wird argumentiert, dass im Fall technikbezogener Konzeptübertragungen die Annahme einer Problemähnlichkeit der aufeinander bezogenen Bereiche in ähnlicher Weise konstitutiv ist wie dort die der Strukturgleichheit: die Annahme nämlich, dass sich eine bereits bewährte Problemlösung für Probleme eines neuen Technikfeldes deshalb als fruchtbar erweisen könnte, weil sich die jeweils zu lösenden Probleme in relevanten Hinsichten ähneln. Ein Rekurs auf Strukturgleichheit, so wird gezeigt, kann die Analogiebeziehung im Fall technikbezogener Konzeptübertragungen dagegen nicht fundieren. Dass entsprechende Annahmen über Strukturgleichheit unter Umständen dennoch getroffen werden, ist eine scheinbare Widersprüchlichkeit, für die der Beitrag eine Antwort sucht. Die konzeptionellen Überlegungen werden anhand einer empirischen Studie über Konzeptübertragungen in der Multiagentensystem-Forschung exemplarisch veranschaulicht.

Einleitung

Erfindungen sind, so S. Colum Gilfillan (1935/1970: 6), in der Regel neue Kombinationen von Elementen des bestehenden Standes von Wissenschaft und Technik. So ist beispielsweise die Erfindung der Schiffsschraube nichts anderes als das Resultat der Übertragung eines gut bekannten Konstruktionsprinzips auf die Problemstellung des motorisierten Schiffsantriebs, nämlich des Prinzips schräggestellter Schaufeln, das zuvor bereits bei anderen Techniken, der Windmühle etwa (oder noch früher: der Technik des Ruderns), erfolgreich zur Anwendung gekommen war (vgl. Gilfillan 1935/1970: 27f.). Dieser Erklärungsansatz von Gilfillan, einem der Pioniere der soziologischen Innovationsforschung, ist in der neueren Wissenschafts- und Techniksoziologie weitgehend in den Hintergrund getreten. Hierfür scheinen zwei Theorieentwicklungen in besonderer Weise verantwortlich zu sein: das Paradigmen-Konzept Kuhns mit seiner Betonung der im Paradigmenwechsel verkörperten radikalen Diskontinuität der Ereignisse und der sozialkonstruktivistische Ansatz in der Wissenschafts- und Technikforschung mit seiner Abkehr von „ideengeschichtlichen“ Erklärungsmustern.

Im Zuge der Einführung des Paradigmenkonzepts in die technikbezogene Innovationsforschung (vgl. Dosi 1982) kommt es zu der folgenreichen Unterscheidung technischer Neue-

1. Für Kommentare und Anregungen zu einer früheren Fassung dieses Aufsatzes danke ich Thomas Malsch, Martin Meister, Werner Rammert, Jörg Strübing und den Herausgebern und Gutachtern der ZfS. Mein Dank für die Diskussion verschiedener Vortragsversionen der hier präsentierten Überlegungen gilt den TeilnehmerInnen des Forschungskolloquiums „Technik- und Innovationsforschung“ am Institut für Soziologie der TU Berlin und den TeilnehmerInnen der Tagung „Transfer von Modellen zwischen Wissenschaftsgebieten“ der Gesellschaft für Wissenschafts- und Technikforschung (GWTF e.V.).

rungen in radikale und inkrementelle Innovationen. Inkrementelle Innovationen sind demnach technische Weiterentwicklungen entlang von Entwicklungspfaden, deren Richtung vorgezeichnet ist durch ein technologisches Paradigma, d.h. durch eine Anzahl etablierter Konstruktionsprinzipien sowie der auf dieser Grundlage bereits erfolgreich realisierten technischen Arrangements (vgl. Dosi 1982: 152ff.). Radikale Innovationen dagegen sind technische Neuerungen, die gegebene Entwicklungspfade verlassen und sich stattdessen als Versuche der Umsetzung neuer technologischer Paradigmen formieren (vgl. Dosi 1982: 154, 158ff.; Freeman/Perez 1988: 46). Mit dieser Unterscheidung wird sämtliche Kontinuitätlichkeit technischer Entwicklung auf Seiten der Weiterentwicklung bestehender Produktlinien verortet, während die technische Innovation im eigentlichen Sinne sich als diskontinuierlicher Prozess darstellt.

Aber auch der sozialkonstruktivistische Ansatz hat zur Entthematisierung kontinuierkeitsbezogener Ansätze in der Innovationsforschung beigetragen. Hier ist es die Fokussierung auf die je aktuellen Prozesse der sozialen Aushandlung von Forschungsergebnissen unter der Bedingung interpretativer Flexibilität, die dazu führt, dass bereits etablierte Wissensbestände lediglich als Rahmenbedingungen entsprechender Aushandlungsprozesse in den Blick kommen, nicht aber als eigenständige Erklärungsfaktoren. So richtet beispielsweise die Wissenschaftssoziologie Collins'scher Prägung ihr Augenmerk exklusiv auf Situationen so genannter wissenschaftlicher Kontroversen, die gerade dadurch charakterisiert sind, dass das bereits konsentiert Wissen mit Blick auf die strittigen Forschungsfragen keine der konkurrierenden Antwortversuche präferiert (vgl. Collins 1983: 94). Ähnliches gilt für die techniksoziologische Adaption dieser zunächst wissenschaftssoziologischen Überlegungen (vgl. Pinch/Bijker 1984).

Allerdings ist in den letzten Jahren auch innerhalb des sozialkonstruktivistischen Ansatzes erkannt worden, dass man den Spielraum möglicher Aushandlungsprozesse anfangs wohl über- und den Einfluss bereits ausgehandelter Bedeutungsfestlegungen auf Innovationsprozesse (den erreichten Stand von Wissenschaft und Technik also) dagegen unterschätzt habe (vgl. Bijker 1995: 281ff.; Schulz-Schaeffer 2000a: 280ff.) Weiterhin dürfte sich herausstellen, dass kaum eine der neuen Techniken, mit denen sich die soziologische Innovationsforschung befasst hat, im strengen Sinne als radikale Innovation gelten kann – ganz abgesehen von der Schwierigkeit, technologische Paradigmenwechsel überhaupt einigermaßen trennscharf zu identifizieren.² Natürlich sind diese wenigen Bemerkungen nicht geeignet, die Verdienste der angesprochenen Ansätze in Frage zu stellen. Vielmehr geht es mir hier darum, darauf hinzuweisen, dass die Abkehr von kontinuierkeitsbezogenen Erklärungsmustern in der soziologischen Innovationsforschung keineswegs durch die Konzeptionen gedeckt ist, die diese Entwicklung maßgeblich beeinflusst haben. Vor diesem Hintergrund scheint es sinnvoll und notwendig, bei der Analyse von Innovationsprozessen – und nicht nur solcher, die die Verbesserung bestehender Produkte und Prozesse zum Gegenstand haben – zumindest ergänzend wieder stärker nach Erklärungen zu suchen, die den Rekurs des Neuen auf Bekanntes thematisieren. Dazu ist der vorliegende Aufsatz ein Beitrag.

Der Gegenstand, anhand dessen die Erzeugung des Neuen durch Bezugnahme auf Bekanntes im Folgenden exemplarisch untersucht wird, ist das Forschungsfeld der Verteilten Künstli-

2. Was nicht zuletzt damit zusammenhängt, dass technische Innovationen in der Regel auf einer Vielzahl heterogener Wissensbestände aufbauen und eine grundlegende Umorientierung mit Blick auf alle erforderlichen Teilbeiträge einer Innovation deshalb sehr unwahrscheinlich ist. Dies ist im Kern bereits Gilfillans (1935/1970: 3ff.) Argument gegen das Konzept radikaler oder, wie es bei ihm heißt, revolutionärer Innovationen.

chen Intelligenz (VKI). Wir stoßen hier auf eine besonders unerwartete Form des Rückgriffs auf vorgängige Wissensbestände bei der Entwicklung technischer Neuerungen, die uns die Eigenschaften entsprechender Übertragungsprozesse dadurch vielleicht deutlicher als anderenorts vor Augen führt: den Versuch, Software-Systeme nach dem Vorbild sozialer Zusammenhänge zu konstruieren. Mit überraschender Selbstverständlichkeit heißt es in einer neueren maßgeblichen Einführung in das noch junge Forschungsfeld, es bestehe zwischen den beteiligten Forschern ein breiter Konsens, dass „VKI-Systeme von Natur aus sozial sind“³ (Moulin/Chaib-draa 1996: 44). Dies könne man sich durchaus analog zu menschlicher Sozialität vorstellen: So wie menschliche Sozialverbände das Resultat erfolgreicher Handlungskoordination seien, so beschäftige sich auch die Verteilte Künstliche Intelligenz „mit Situationen, in denen verschiedene Systeme interagieren, um ein gemeinsames Problem zu lösen“ (Moulin/Chaib-draa 1996: 4).

Im Folgenden wird es im ersten Abschnitt zunächst darum gehen, ein idealtypisches Konzept technikbezogener Konzeptübertragungen zu entwickeln, das Gilfillans These der Erfindung als neuer Kombination etablierten Wissens und Könnens präziser fasst und insbesondere genauer anzugeben in der Lage ist, unter welchen Bedingungen mit entsprechenden Übertragungsprozessen überhaupt zu rechnen ist. Zu diesem Zweck greife ich auf Max Blacks Konzept des theoretischen Modells zurück. Black zufolge gründen wissenschaftliche Konzepttransfers auf der Annahme einer Strukturgleichheit zwischen den Phänomenen zweier Forschungsbereiche, die es erlaubt, etabliertes Wissen aus dem einen Forschungsfeld als theoretische Innovation in das andere Forschungsfeld zu importieren. Ich werde argumentieren, dass im Fall technikbezogener Konzeptübertragungen in analoger Weise die Annahme einer Problemähnlichkeit der aufeinander bezogenen Bereiche konstitutiv ist. Die drei daran anschließenden Abschnitte dienen dazu, diese Konzeption am Beispiel der Multiagentensystem-Forschung der VKI zu überprüfen und weiterzuentwickeln. Im zweiten Abschnitt wird das Problem der Koordination eigenständig verhaltensfähiger Entitäten als das Problem identifiziert, mittels dessen in der Multiagentensystem-Forschung die Ähnlichkeitsbeziehung zur Welt des Sozialen hergestellt wird. Der dritte Abschnitt diskutiert Bestrebungen der Verteilten Künstlichen Intelligenz, diese Problemähnlichkeit zu nutzen, um Mechanismen menschlicher Verhaltensabstimmung als technische Problemlösungen für das Problem der Agentenkoordination zu reformulieren. Dabei wird sich zeigen, dass dieses Vorgehen zwar einige Erfolge aufzuweisen hat, dass die formulierte Problemähnlichkeit über das Koordinationsproblem insgesamt aber recht unspezifisch ist und deshalb zu recht unspezifischen Analogieschlüssen führt. Als Konsequenz daraus erfolgt eine nach Maßgabe des bis dahin Gesagten überraschende Wende: Eine Umorientierung auf Strukturgleichheit als Grundlage der Ähnlichkeitsbeziehung. Sie ist das Thema des vierten Abschnitts. Überraschend ist sie, weil mit der Umorientierung von Problemähnlichkeit auf Strukturgleichheit das zentrale Kalkül dispensiert wird, das den Rekurs auf Bekanntes in der Technikentwicklung rechtfertigt: Dass sich der Rückgriff auf eine bekannte Problemlösung lohnen könne, weil das anstehende Problem dem dortigen Problem hinreichend ähnlich ist. Es lässt sich allerdings begründen, dass der Umweg über Strukturgleichheit im Wesentlichen eine indirekte Form ist, Problemähnlichkeit herzustellen. Im letzten Abschnitt wird das Konzept technikbezogener Konzeptübertragung in entsprechender Weise erweitert.

3. Hier wie bei allen folgenden im Original englischsprachigen Zitaten stammen die Übersetzungen vom Verfasser.

1. Wissenschaftliche und technikbezogene Konzeptübertragungen: Strukturgleichheit vs. Problemähnlichkeit

Ausgangspunkt der folgenden Überlegungen ist die These der innovatorischen Bedeutung einer Vorgehensweise, die in der Wissenschaftsphilosophie seit Jahrzehnten intensiv diskutiert wird (vgl. Radman 1995), die in der Wissenschaftssoziologie nach ungnädiger Kommentierung durch Knorr-Cetina (1984: 92ff.) kaum weitere Beachtung gefunden hat,⁴ die dafür aber in der Techniksoziologie im Zusammenhang mit dem leider reichlich diffusen Begriff des Leitbildes in den letzten Jahren einige Aufmerksamkeit für sich verbuchen konnte (vgl. Mambrey et al. 1995). Die einschlägige Beschreibung, wie man sich diese Vorgehensweise idealtypisch vorzustellen habe, stammt nach wie vor von Max Black, dessen Überlegungen zugleich den Startpunkt der wissenschaftsphilosophischen Diskussion bildeten. Sie lautet: „1. Wir haben ein primäres Untersuchungsfeld, innerhalb dessen *einige* Tatsachen und Regelmäßigkeiten bereits etabliert sind ... 2. Es gibt einen Bedarf ... nach weiterer wissenschaftlicher Beherrschung des primären Bereichs. 3. Wir beschreiben einige Entitäten (Objekte, Materialien, Mechanismen, Systeme, Strukturen), die zu einem relativ unproblematischen, bekannteren oder besser organisierten zweiten Bereich gehören. Die postulierten Eigenschaften dieser Entitäten werden in der jeweils als nützlich erachteten Detailliertheit beschrieben. 4. Es sind implizite oder explizite Korrelationsregeln verfügbar, die es erlauben, Aussagen über das zweite Feld in korrespondierende Aussagen über das primäre Feld zu übersetzen. 5. Folgerungen aus den Annahmen, die in dem zweiten Feld aufgestellt worden sind, werden mittels der Korrelationsregeln übersetzt und dann unabhängig überprüft an bekannten oder vorhergesagten Daten aus dem ersten Bereich.“ (Black 1962: 230) Bedingung für die Entdeckung solcher Korrelationsregeln ist, dass die aus dem Referenzbereich stammenden Konzepte die gleichen Strukturen bzw. die gleichen Muster von Beziehungen beschreiben, die auch in dem eigentlich interessierenden Feld von Bedeutung sind: „(D)er Schlüssel zum Verständnis der gesamten Transaktion ist die Strukturgleichheit, die es in günstigen Fällen erlaubt, aus Annahmen über den zweiten Bereich Einsichten über das ursprüngliche Feld des Interesses zu gewinnen“ (Black 1962: 230f., vgl. auch ebd.: 222f.).

Konzeptübertragungen dieser Art bezeichnet Black als theoretische Modelle.⁵ Ihre Verwendung ähnele der Verwendung von Metaphern. So wie starke Metaphern⁶ strukturelle Korrespondenzen vermittelten und auf diese Weise unter Umständen neue Erkenntnis hervorbringen könnten (vgl. Black 1977/1983: 396, 404ff.), so zielten auch theoretische Modelle auf die Möglichkeit, neue Verbindungen zu sehen, wobei der Innovationserfolg im Vorhinein natür-

4. Knorr-Cetina (1984: 116ff., 123f.) wirft den gleich anzusprechenden Überlegungen vor, den beobachteten Innovationserfolg retrospektiv zur Erklärung des Innovationsgeschehens zu verwenden (dies ist das gängige sozialkonstruktivistische Argument gegen „ideengeschichtliche“ Erklärungen, ähnlich z.B. auch Pinch/Bijker 1984: 405f.) und dabei die Fabrikations- und Verhandlungsprozesse in der alltäglichen Praxis des Labors, die diesen Erfolg tatsächlich bewirkten, zu ignorieren. In einer neueren Veröffentlichung gelangt sie dagegen zu der deutlich positiveren Einschätzung, die Interpretation von Metaphern und Analogien in der Tradition Blacks erweise sich „als wichtig für das Verständnis konzeptueller Innovationen in Wissenschaft und anderen Bereichen“ (Knorr-Cetina 1995: 333).

5. Der Begriff des Modells bezeichnet hier also das Ergebnis des Übertragungsprozesses. Wegen dieser terminologischen Festlegung enthielte die Bezeichnung der Übertragungsprozesse als Modelltransfers eine unschöne Dopplung des gemeinten Sachverhalts. Ich verwende deshalb stattdessen den Begriff der Konzeptübertragung oder den des Konzepttransfers.

6. Vgl. Black 1977/1983: 389ff. Angesprochen ist damit die Metapher, die im Sinne der von Black (1954/1983: 68ff.; 1977/1983: 391ff.) propagierten Interaktionstheorie der Metapher wirksam wird.

lich nicht garantiert werden kann (vgl. Black 1962: 237). Diese Nähe zu der unterstellten innovativen Wirksamkeit starker Metaphern hat den hier skizzierten Überlegungen die Bezeichnung als „Metapher-Theorie der Innovation“ (Knorr-Cetina 1984: 92) eingetragen.⁷ Der Begriff der Metapher, dies zu betonen ist Black wichtig, sollte jedoch nicht dem Missverständnis Vorschub leisten, dass Wissenschaftler, die solche Konzeptübertragungen vornehmen, sich lediglich von diffusen, nicht weiter begründbaren Ahnungen leiten ließen. Vielmehr gibt es „eine rationale Grundlage“ der Verwendung theoretischer Modelle, nämlich die „mutmaßliche Isomorphie zwischen Modell und Anwendungsfeld“ (Black 1962: 238).

Black's Paradebeispiel für theoretische Modelle ist Clerk Maxwells Konzept des elektromagnetischen Feldes (vgl. Black 1962: 226ff.). Ich greife hier auf die vereinfachende Darstellung bei Miller (1995: 201) zurück, die sich für illustrative Zwecke besonders eignet. Demnach beruhen Maxwells Überlegungen zum Elektromagnetismus auf der Unterstellung der folgenden Isomorphie: „Das elektromagnetische Feld verhält sich *als ob* es aus einer Ansammlung von Gummibändern und Umlenkrollen bestünde.“ (Miller 1995: 201) Maxwell stellt damit eine Strukturanalogie zu einem wissenschaftlich gut beherrschten Referenzbereich her, nämlich „zu der gut verstandenen klassischen Mechanik von Gummibändern und Umlenkrollen mit ihrer Mathematik, ihren physikalischen Konzepten und ihrer Anschaulichkeit“ (Miller 1995: 201). Das konzeptuelle Wissen über die Strukturen und Beziehungsmuster des dortigen Forschungsgegenstandes wird mittels der Analogie aufgegriffen und zum Zweck der Erklärung des noch unzureichend verstandenen Begriffs des elektromagnetischen Feldes adaptiert.

Die Forschungsstrategie ist in ihrer idealisierten Form mithin recht einfach: Man beobachtet in einem Forschungsfeld bestimmte erklärungsbedürftige Regelmäßigkeiten. Im günstigen Fall stößt man bei der Suche nach einer Erklärung auf ein anderes Forschungsfeld, in dem vergleichbare Regelmäßigkeiten beobachtet worden sind und in dem es zudem eine etablierte und gründlich erforschte Theorie zu deren Erklärung gibt. Diese Theorie würde mit anderen Worten, ließe sie sich auf das interessierende Forschungsfeld übertragen, eine konsistente Erklärung der fraglichen Regelmäßigkeiten bieten. Eine solche Übertragung wird deshalb versucht vorzunehmen.

Nicht nur in der wissenschaftlichen Forschung, auch im Kontext technischer Innovationen stoßen wir auf Konzeptübertragungen dieser Art. Ein bekanntes Beispiel ist das Schreibklavier, eine frühe und wirkungsgeschichtlich außerordentlich einflussreiche Form der mechanischen Schreibmaschine, deren Konstruktion die Hebelmechanik des Klaviers adaptiert, um nun Buchstaben statt Töne anzuschlagen (vgl. Knie 1989: 12ff.; Mambrey et al. 1995: 74ff.). Allerdings erfolgen Konzeptübertragungen hier unter etwas anderen Bedingungen als dort. Der hauptsächliche Unterschied ist der, dass das primäre wissenschaftliche Erkenntnisinteresse an Konzeptübertragungen in der Erklärung empirisch beobachtbarer Phänomene der vorgefundenen natürlichen oder sozialen Welt besteht – auch wenn diese zu Erkenntniszwecken wie immer gesäubert, experimentell zugerichtet und theoretisch vorstrukturiert wird. Im Kontext technikbezogener Forschung dagegen dienen Konzeptübertragungen nicht der Erklärung eines Gegenstandes, sondern seiner Herstellung. Diese Differenz hat eine wichtige Konsequenz: Sich andeutende Strukturgleichheit zwischen dem Gegenstand des Interesses und bestimmten Phänomenen eines anderen Forschungsfeldes kann im Fall technikbezogener Kon-

7. Auch wenn Black zu Recht als einer der Autoren rezipiert wird, die den Begriff der Metapher für die Wissenschaftsphilosophie fruchtbar gemacht haben (vgl. Danneberg et al. 1995: 13ff.), ist – wenigstens in seinen späteren Veröffentlichungen – der Begriff des Modells der grundlegendere. Dies drückt sich in seiner These aus, jede Metapher sei „die Spitze eines untergetauchten Modells.“ (Black 1977/1983: 396)

zeptübertragungen nicht als Kriterium der Formulierung von Analogierelationen erhalten. Denn die Struktur des Forschungsgegenstandes, also ein bestimmter technischer Wirkungszusammenhang, ist ja genau das, was erst als Resultat des Forschungsprozesses entsteht.

Dennoch scheint auch in der technologischen Forschung und der Technikentwicklung mitunter beträchtliche Zuversicht zu herrschen, dass man mit einer bestimmten Form des Konzepttransfers auf Erfolg versprechenden Pfaden wandelt. Eine solche Zuversicht kommt etwa in der oben zitierten Überzeugung zum Ausdruck, man müsse die Multiagentensysteme der VKI als von Natur aus sozial betrachten. Das nährt die Vermutung, dass es auch hier ein einschlägiges Kriterium gibt. Im Fall technikbezogener Konzeptübertragungen, so meine These, bildet eine sich andeutende Problemähnlichkeit die rationale Grundlage für die Herstellung von Analogiebeziehungen. Dies lässt sich wiederum am Beispiel des Schreibklaviers gut illustrieren: Die Bezugnahme auf Konstruktionswissen aus dem Klavierbau kommt dadurch zu Stande, dass eine Ähnlichkeitsrelation hergestellt wird zwischen dem Problem des mechanisch unterstützten Schreibens und dem Problem, Saiten mechanisch anzuschlagen. Auf der Grundlage dieser Ähnlichkeitsbeziehung erscheint es dann als viel versprechend, die im Klavierbau bewährte technische Lösung dieses Problems in geeignet modifizierter Form zu verwenden, um auf diese Weise nun auch das Problem mechanischen Schreibens zu lösen.

Da es zwischen den beiden Feldern einer technikbezogenen Konzeptübertragung nie eine vollständige Problemgleichheit geben kann – anderenfalls hätte man das Problem ja bereits gelöst – ist auch in diesem Fall der Erfolg keineswegs garantiert. Es kann stets nur ein gewisses Maß an Problemähnlichkeit geben und dies impliziert zugleich, dass die aufeinander bezogenen Probleme in bestimmten Aspekten durchaus unähnlich sind. Es kann also nicht von vornherein ausgeschlossen werden, dass die Adaption einer technischen Problemlösung in der neuen Problemumgebung nicht doch in entscheidenden Hinsichten versagt. Der Fall des Schreibklaviers illustriert, dass mit solchen Schwierigkeiten selbst bei insgesamt erfolgreichen Konzepttransfers zu rechnen ist. So erweist sich beispielsweise die auf die Schreibmaschine übertragene Hebelmechanik des Klaviers als ungeeignet, das Problem gleichzeitigen Anschlagens von Tasten und daraus resultierenden Verhakens von Typenhebeln zu vermeiden. Dieses Problem hat die mechanische Schreibmaschine als ungelöstes mit ins Grab genommen. Die Hebelmechanik des Klaviers sieht dafür schlichtweg keine Lösung vor. Denn beim Klavier spielen, ist die Möglichkeit des gleichzeitigen Anschlagens mehrerer Töne ja durchaus erwünscht.

Mit den voranstehenden Überlegungen habe ich selbst eine Konzeptübertragung vorgenommen. Ich habe argumentiert, dass die zumindest innerhalb der Wissenschaftsphilosophie gut etablierte Theorie wissenschaftlicher Konzeptübertragungen in geeignet modifizierter Form ein vielversprechendes theoretisches Modell zur Erklärung technikbezogener Konzeptübertragungen ist. Die wesentlichen Gesichtspunkte dieses Erklärungsmodells sind, um es kurz zusammenzufassen: (1) Es wird eine sich abzeichnende Problemähnlichkeit zwischen dem interessierenden technischen Problem und der Problemstellung eines beliebigen anderen Technikfeldes festgestellt. (2) In diesem anderen Technikfeld gibt es für die dortigen Problemstellungen bewährte technische Lösungen, also ein entsprechendes Regel- und Konstruktionswissen samt dessen beispielhafter Verkörperung in funktionierenden technischen Abläufen. (3) Es wird versucht, diese technischen Problemlösungen in einer Weise zu modifizieren, die es erlaubt, sie auf das interessierende Problem zu übertragen.⁸

8. Gegen die hier vorgestellte Konzeption könnte der grundlegende Einwand erhoben werden, sie beruhe auf einer Unterscheidung zwischen wissenschaftlicher und technologischer/technischer Forschung, die in dieser

Nun muss man allerdings davon ausgehen, dass technikbezogene Konzeptübertragungen im Regelfall nicht so einfach und klar strukturiert sind, wie es das bislang zur Illustration verwendete Beispiel des Schreibklaviers nahe legt. So wird man erstens konzедieren müssen, dass in den Frühphasen technischer Innovationen häufig nicht sehr präzise angegeben werden kann, worin die Probleme bestehen, die eine projektierte technische Lösung bewältigen soll. Dies ist insbesondere dann der Fall, wenn es nicht um technische, sondern um technologische Probleme geht, also um Problemstellungen, die auf ein übergreifendes Lösungsmuster für eine Menge mehr oder weniger gleichgearteter Probleme zielen. Daraus folgt zweitens, dass eine aufgefundene Problemähnlichkeit dann anfangs notwendigerweise auf wackeligen Beinen steht. Unter solchen Bedingungen ist damit zu rechnen, dass der Rekurs auf ein anderes Forschungsfeld nicht nur bei der Problemlösung eine Rolle spielt, sondern bereits bei der Problemformulierung. Drittens schließlich wird bei weitem nicht immer die günstige Situation eintreten, dass in dem anderen Bereich bereits für alle potenziell relevanten dortigen Problemstellungen gut durchgearbeitetes und bewährtes technisches Wissen vorliegt. Für das Forschungsfeld, das im Folgenden näher betrachtet werden soll, die Multiagentensystem-Forschung der VKI, trifft jede dieser drei Bedingungen zu.

2. Das Koordinationsproblem als Grundlage der Ähnlichkeitsbeziehung zwischen Multiagentensystemen und der Welt des Sozialen

Die technischen Artefakte, um deren Entwicklung es der Verteilten Künstlichen Intelligenz geht, Multiagentensysteme, lassen sich beschreiben als Systeme wie immer aufeinander abgestimmten Operierens einer Mehrzahl so genannter Agenten. Als Agenten werden Software-Module bezeichnet, die über bestimmte, von ihnen selbst gesteuerte Aktionsprogramme verfügen und in der Lage sind, ihre eigenen Aktionen unter Berücksichtigung derer anderer Agenten (bzw. allgemeiner: unter Berücksichtigung bestimmter Gegebenheiten ihrer Umwelt) selbstständig auszuwählen (Bond/Gasser 1988). Konstitutiv für die Multiagentensystem-Forschung sind die folgenden drei Merkmale: (1) *Verteiltheit*: Gegenstand ist die Entwicklung von Problemlösungsmechanismen, in denen die Bearbeitung des fraglichen Problems auf mehrere Agenten verteilt ist; (2) *Autonomie*: Die einzelnen Agenten werden als (teil)autonome und (teil)intelligente Einheiten konzipiert, die sich in ihrem Verhalten nach eigenen Aktionsprogrammen richten und dabei über ein gewisses Maß eigenständiger Problemlösungskompetenz verfügen; (3) *Koordination*: Die Problemlösung erfolgt nicht nach Maßgabe eines vom Programmierer implementierten Ablaufs, sondern resultiert aus der Koordination zwischen den Agenten.

Von den beteiligten Wissenschaftlern selbst wird das Forschungsfeld der Verteilten Künstlichen Intelligenz in der Regel als ein Teilgebiet der Künstliche Intelligenz-Forschung (KI)

Trennschärfe empirisch zunehmend weniger vorzufinden sei. In der Tat lässt sich kaum bestreiten, dass eine Vielzahl technologischer/technischer Entwicklungen auf wissenschaftliche Forschung aufsetzt (Stichwort: neues generisches Wissen, vgl. Nelson 1988: 314), so wie umgekehrt Technik in vielen Fällen zu einer unverzichtbaren Voraussetzung wissenschaftlicher Forschung geworden ist. Es wäre dennoch vorschnell, nun von einem nicht mehr weiter differenzierbaren Komplex der „technoscience“ (Latour 1987: 174f.) auszugehen und damit zugleich auch den analytischen Wert der Unterscheidung zu leugnen. Denn es macht nach wie vor einen wesentlichen Unterschied für das je beobachtete Forschungshandeln, ob es sich am Ziel der Erklärung empirischer Beobachtungen orientiert oder am Ziel der Konstruktion funktionsfähiger Abläufe. Auf diese Differenz kommt es im vorliegenden Zusammenhang an. Um ihre willens nehmen ich in Kauf, dass die Unterscheidung etliche Forschungsgebiete ganz oder teilweise als technologische Forschungsfelder markiert, die ihrer institutionellen Ansiedlung nach als Wissenschaften gelten (z.B. Ingenieur-, Bio- oder Erziehungswissenschaften) und dass sie andere (Mathematik, Logik und Philosophie) überhaupt nicht einzuordnen erlaubt.

eingeeordnet, das über die klassische KI allerdings in einer Hinsicht hinausgeht, indem es nicht das einzelne intelligente System zum Gegenstand hat, sondern eine Mehrzahl zusammenwirkender intelligenter Einheiten. Eine Äußerung für viele: „VKI kann man als Unterfeld der KI betrachten. Dass du sagst: KI allgemein: intelligente Systeme zu bauen, effiziente, intelligente Systeme, dem Menschen irgendwie abgeguckt. Und dann zu sagen: Okay, die VKI speziell verwendet diesen Ansatz, kleine intelligente Einheiten zu bauen, durch deren Kooperation zusammen noch irgendwas Größeres entsteht. Also insofern ist es ein Unterfeld, es benutzt eben dieselben Paradigmen, um Agenten zu bauen, aber komplementär, indem du sagst: Okay, da kommt noch so eine Komponente dazu, so was wie Emergenz beispielsweise.“ (FW3, 8: 21-27)

Die Verteilte Künstliche Intelligenz ist ein vergleichsweise junges Forschungsgebiet. Die Forschungen, die heute als die Pionierarbeiten gelten, wurden Ende der Siebziger- bis Mitte der Achtzigerjahre in den USA durchgeführt. In Deutschland ist die VKI-Forschung gerade einmal ein gutes Jahrzehnt alt. Ende 1989 fand am Rande des German Workshop of Artificial Intelligence ein erstes Treffen einer Hand voll Informatiker statt, mit dem Ziel, das gemeinsame Interesse an den neuen Forschungsideen zu bündeln. Ihr Engagement mündete 1993 in die Gründung der Fachgruppe „Verteilte Künstliche Intelligenz“ in der Gesellschaft für Informatik. Der deutschen VKI-Forschung ist es dann offenbar recht schnell gelungen, Anschluss an die internationale (und das heißt in diesem Fall primär: die u.s.amerikanische) Forschung zu gewinnen. Dafür jedenfalls spricht die Präsenz deutscher Wissenschaftler bei den einschlägigen Tagungen und in den einschlägigen Veröffentlichungen.

Das junge Alter der Verteilten Künstlichen Intelligenz und die erst beginnende wissenschaftliche Institutionalisierung spiegelt sich in der Zusammensetzung der beteiligten Wissenschaftler wider. So gab es nach Einschätzung eines Gesprächspartners (UP3, mündliche Auskunft) Ende der Neunzigerjahre in Deutschland nur zwei oder drei Universitätsprofessoren, die von ihren Forschungsschwerpunkten her als VKIler bezeichnet werden können.⁹ Das Gros der Forschungsarbeiten stammt aus der Gruppe der wissenschaftlichen Mitarbeiter, die, um eine ungefähre Größenordnung anzugeben, Ende der Neunzigerjahre etwa 30 bis 50 Wissenschaftler umfasste. Auch die Initiative zur Etablierung der Verteilten Künstlichen Intelligenz in Deutschland wurde zunächst in der Hauptsache von Forschern aus dem wissenschaftlichen Mittelbau betrieben. Diese Wissenschaftler der ersten Stunde sind heute jeweils etwa zur Hälfte Lehrstuhlinhaber oder in zumeist außeruniversitären Forschungszusammenhängen als Projekt- oder Forschungsgruppenleiter tätig. Die meisten der von uns Befragten sind über die Beschäftigung mit Problemen der KI-Forschung zur VKI gelangt. „Genuine“ VKI-Forscher, also solche, die sich bereits im Rahmen ihrer akademischen Ausbildung auf dieses Forschungsfeld spezialisiert haben, finden sich nur in der Generation der jüngeren Nachwuchswissenschaftler.

9. Eigene Recherchen ergeben für diesen Zeitpunkt etwa die doppelte Zahl. Solche Einschätzungen hängen jedoch stark von den jeweils vorausgesetzten Grenzziehungen ab, die in dem in Entwicklung begriffenen Forschungsfeld noch nicht sehr ausgeprägt objektiviert sind. Es kann hier deshalb nur um die Angabe einer ungefähren Größenordnung gehen. Die entsprechenden Grenzziehungsprobleme werden auch von den Befragten selbst ausführlich thematisiert: „Es ist sehr schwer auch für Leute, die meinetwegen neu in das Gebiet rein wollen, zu sehen, was ist denn hier eigentlich los? Also das Gebiet ist sehr sehr unklar strukturiert. Das ist eines der Probleme.“ (IW4, 23: 13-17) Insbesondere wird beklagt, dass, nachdem Agenten zu einem Modethema der Informatik geworden sind, in den letzten Jahren „sehr sehr viele Leute auf diesen Agenten-Wagen drauf springen“ (IW4, 23: 10) und den Begriff aus eher forschungstaktischen denn inhaltlichen Gründen für ihre Forschung beanspruchen.

Diese Zusammensetzung der Wissenschaftler findet sich auch in dem Sample der Befragung wieder, die meine KollegInnen und ich im Rahmen des DFG-Projekts „Sozialmetaphern in der Verteilten Künstlichen Intelligenz“ durchgeführt haben.¹⁰ Das Sample besteht aus drei Universitätsprofessoren (UP1-3), elf universitär beschäftigten wissenschaftlichen Mitarbeitern (UW1-11), sechs Wissenschaftlern an außeruniversitären akademischen Forschungseinrichtungen (FW1-6) und sechs wissenschaftlichen Mitarbeitern in industriellen Forschungseinrichtungen, davon einer in leitender Position (IW1-5, IL1). Es umfasst mit zwei Ausnahmen alle Wissenschaftler, die von den Beteiligten selbst als die Pioniere der Verteilten Künstlichen Intelligenz in Deutschland genannt wurden. Die Interviews wurden als leitfadengestützte Experteninterviews von jeweils etwa anderthalb bis zwei Stunden Länge durchgeführt und anschließend transkribiert. Nach einer explorativen Phase Ende 1997 wurde der Interviewleitfaden überarbeitet und im Frühjahr und Sommer 1998 der Hauptteil der Interviews durchgeführt. Dem Forschungsinteresse des Projektes entsprechend, waren die Interviewfragen im Kern darauf gerichtet herauszufinden, ob und in welcher Weise außerinformatische, und hier insbesondere: auf Sozialverhalten abzielende Konzepte und Vorstellungen in den technologischen Innovationsprozess der VKI einfließen. Die transkribierten Interviews wurden anschließend nach einem an Strauss (1987) orientierten Verfahren der Kategorienbildung vergleichend ausgewertet. Die so gewonnenen Ergebnisse wurden in einer Vielzahl späterer informeller Nachfragen bestätigt bzw. verfeinert.

Hinsichtlich ihrer allgemeinen Forschungsorientierung unterscheidet sich (nicht nur) die deutsche VKI in einem Punkt deutlich von ihrer Mutterdisziplin: Während sich die KI in eine kognitionswissenschaftliche und eine anwendungsorientierte Richtung ausdifferenziert hat, die sich zum Teil in einer Weise auseinanderentwickelt haben, dass die Konzepte des theoretischen Zweigs für die konkret problemorientierte Forschung als hinderlich angesehen werden (vgl. Schlese 1995: 376ff.), ergibt sich für die Verteilte KI ein sehr viel einheitlicheres Bild. Sie organisiert sich ganz überwiegend als problemlösungsorientierte technologische Forschung. Die Frage nach dem „Wesen“ der menschlichen Intelligenz und den grundsätzlichen Möglichkeiten und Grenzen ihrer technischen Nachbildung, die der kognitionswissenschaftlichen KI zu Grunde liegt, tritt in der Verteilten Künstlichen Intelligenz in den Hintergrund gegenüber der sehr viel anwendungsnäheren Frage danach, wie man eine Anzahl von Agenten in einem Multiagentensystem dazu bringt, gemeinsam ein Problem zu lösen.

Das Problem, dessen Lösung man sich von Multiagentensystemen verspricht, ist ein technologisches Problem im bereits angesprochenen Sinne. Es geht darum, technische Verfahren zur Bewältigung eines bestimmten Typs von Problemen zu entwickeln, die als „verteilte Probleme“ bezeichnet werden. Die zu Grunde liegende Problemwahrnehmung lautet: „Bei vielen Problembereichen liegt bereits eine natürlich gegebene Verteilung vor und legt daher eine Anwendung der VKI nahe.“ (Martial 1992b: 8) Dies gelte etwa für räumliche oder funktionale Verteilung (vgl. Davis/Smith 1983: 67). Ein Beispiel für räumliche Verteilung ist das Problem der Interpretation und Integration von Daten an verschiedenen Orten aufgestellter Sensoren, etwa bei Verkehrsüberwachungssystemen (vgl. Lesser/Corkhill 1983). Als Beispiel für funktionale Verteilung werden Probleme angeführt, deren Bearbeitung ein hochgradig ausdifferenziertes Spezialistentum voraussetzt, etwa im Bereich der medizinischen Diagnose (vgl. Durfee et al. 1989: 64). Unter solchen Bedingungen sei es von Vorteil, Teilprobleme des zu

10. Das Projekt wurde unter der Leitung von Thomas Malsch von Holger Braun, Alexandra Engels, Peter Imhof, Rolf Lührs und dem Verfasser durchgeführt. Zu einer ausführlicheren Darstellung der Forschungsergebnisse vgl. Schulz-Schaeffer 2000b.

lösenden Gesamtproblems dort zu lösen, wo sie anfallen, und die Teillösungen dann zu integrieren. Bei den so genannten verteilten Problemen handelt es sich, weniger substantialistisch ausgedrückt, mithin um Probleme, von denen angenommen werden kann, dass die Vorteile einer verteilten Problembearbeitung deren Nachteile überwiegen, wobei die Gründe hierfür recht unterschiedlich sein können. (vgl. Martial 1992a: 16; Durfee et al. 1987: 31f.)

Die Herangehensweise, verteilte Probleme verteilt zu lösen, indem Teillösungen von unterschiedlichen, jeweils für sich problemlösungsfähigen Agenten erzeugt werden, transformiert das jeweilige Einzelproblem (Wie löse ich ein bestimmtes verteiltes Problem befriedigend?) in ein vorgelagertes, und zunächst sehr viel allgemeineres Problem. Es lautet: Wie bringe ich Agenten, die ein gewisses Maß an Verhaltensautonomie besitzen müssen, um Teilprobleme selbstständig lösen zu können, zu einer wie immer gearteten Form koordinierten Verhaltens? Diese Frage benennt das Basisproblem und zentrale Thema der Multiagentensystem-Forschung: das Koordinationsproblem (vgl. Jennings 1996: 187; Van de Velde/Perram 1996: VIII). Es ist zugleich der zentrale Anknüpfungspunkt für die Herstellung der Analogierelation zur Welt des Sozialen. Die Frage, die die Suche nach potenziell nutzbringenden Analogien leitet, lautet nun: Wo findet man außerhalb der Multiagentensystem-Forschung Bereiche, in denen das Problem die Koordination einer Mehrzahl zu selbstständigem Verhalten fähiger Entitäten ist. Und die Antwort, die die Multiagentensystem-Forschung findet, lautet: in der Welt des Sozialen, also im Bereich des zusammenwirkenden Verhaltens oder Handelns von Lebewesen.

3. Problemähnlichkeit: Die Welt des Sozialen als Bezugsquelle technischer Problemlösungen

Gemessen an der oben skizzierten idealtypischen Vorgehensweise ist die Welt des Sozialen kein sonderlich günstiges Bezugsfeld für technikbezogene Konzeptübertragungen. Um das zu sehen, muss man nur das Paradebeispiel des Schreibklaviers zum Vergleich heranziehen: Dort stellt die wahrgenommene Problemähnlichkeit einen Bezug zu einem eng umgrenzten Technikfeld her, in dem es für eine klar definierte Problemstellung (mechanisch unterstütztes Anschlagen von Saiten) eine einzige maßgebliche technische Lösung gibt (die Hebelmechanik) samt eines gut etablierten und weitgehend homogenen professionalisierten Wissens darüber, wie sich diese Lösung erzeugen lässt. Die Welt des Sozialen dagegen präsentiert sich – selbst wenn sich die Analogiebeziehung nur auf menschliche Sozialität erstreckt – in sehr unterschiedlichen wissenschaftlichen Konzepten. Auch liegt ein wesentlicher Teil dieses Weltwissens als Alltagswissen vor, also in nicht gesondert theoretisch durchgearbeiteter Form. Hinzu kommt, dass die Welt des Sozialen zunächst primär kein Technikfeld ist. Das heißt man kann nicht umstandslos davon ausgehen, dass soziale Strukturen und Beziehungsmuster als technische Lösungen für bestimmte Probleme ersonnen und konstruiert worden sind. Die Rede von Problemen und Problemlösungen beruht in vielen Fällen eher auf den Rekonstruktionen eines Beobachters, der bestimmte Muster handelnden Zusammenwirkens als funktional mit Blick auf bestimmte Problemstellungen deutet. Umgekehrt bedarf es einer sehr spezifischen Zurichtung sozialer Arrangements, damit diese auch prospektiv als Techniken funktionieren, also als Zusammenhänge, mit deren Hilfe sich hinreichend zuverlässig und wiederholbar bestimmte erwünschte Effekte erzielen lassen (zu dieser Definition vgl. Schulz-Schaeffer 1999: 410).

Im Gegensatz zu dem Konstruktionswissen über die Herstellung der Hebelmechanik für Klaviere präsentiert sich professionelles oder alltägliches Wissen über die Verhaltenskoordination von Lebewesen nicht ohne weiteres als bewährtes technisches Wissen dergestalt, dass man

auf seiner Grundlage entsprechende Zusammenhänge koordinierten Verhaltens gezielt einrichten könnte. Das Problem der Konzeptübertragung besteht hier also nicht allein in der Umstellung von „Sozialreferenz“ auf „Computerreferenz“ (so auch Malsch 1997: 16f.). Vielmehr muss das importierte Wissen vielfach überhaupt erst als prospektiv nutzbares technisches Problemlösungswissen reformuliert werden. In der Tat lassen sich einige der wichtigsten Koordinationsmechanismen von Multiagentensystemen als in einer solchen Weise reformulierte sozialwissenschaftliche oder alltagstheoretische Konzepte des handelnden Zusammenwirkens von Menschen verstehen. Wie erfolgt diese Rekonstruktion? Zur Beantwortung dieser Frage betrachte ich die beiden entwicklungsgeschichtlich einflussreichsten Koordinationskonzepte der Multiagentensystem-Forschung: die Gutwilligkeitsannahme und das Kontraktnetz-Protokoll.

Die Gutwilligkeitsannahme

In einem Bericht über den ersten Workshop des neu entstehenden Forschungsgebiets der Verteilten KI, der 1980 am MIT in Boston stattfand, hält Davis (1980: 42) fest: „(W)ir betrachten Kooperation im Sinne von gutwilligem Problemlösungsverhalten, d.h. wie können Systeme, die völlig bereitwillig sind, sich wechselseitig aneinander anzupassen, so handeln, dass sie ein erfolgreiches Team bilden?“ Der grundlegende Koordinationsmechanismus, der hier ins Auge gefasst wird, ist die generelle Hilfsbereitschaft zwischen den beteiligten Agenten. Diese werden konzipiert als „freundliche Agenten, die sich wünschen zu tun, worum sie gebeten werden“ (Martial 1992a: 41).

Der Koordinationsmechanismus allgemeiner Hilfsbereitschaft kommt in der Multiagentensystem-Forschung nicht nur bei einfachen Problemen der Verteilung von Aufgaben zum Zuge (vgl. Connah/Wavish 1990). Er wird durchaus auch in komplexeren Zusammenhängen eingesetzt, etwa wenn es darum geht, Teillösungen einer Aufgabe so aufeinander abzustimmen, dass keine der Teillösungen die Realisierung anderer Teillösungen verhindert. Ein Beispiel hierfür ist das von Cammarata et al. beschriebene Szenario für ein agentenbasiertes Flugverkehrsüberwachungssystem. Die den Flugzeugen zugeordneten Agenten sollen in diesem Szenario im Fall drohender Kollisionen zwischen zwei Flugzeugen Ausweichstrategien entwickeln. Der Plan eines einzelnen Agenten zur Modifikation der Flugroute, der sein Flugzeug aus der Gefahrenzone bringen würde, wird dabei nur unter der Bedingung realisiert, dass sich aus ihm nicht weitere potenzielle Kollisionen zwischen Flugzeugen in dem betreffenden Luftraum ergeben, für die keine Lösungen gefunden werden können (vgl. Cammarata et al. 1983/1988; Steeb et al. 1981/1988). Die übergreifende Problemlösung formiert sich dabei als emergentes Produkt eines kooperativen Verhaltensmusters der Beteiligten (vgl. Rammert 1998: 95f.).

Koordination auf der Grundlage hilfsbereiter Agenten ist nach Aussagen eines unserer Gesprächspartner auch in der gegenwärtigen Praxis der Multiagentensystem-Forschung noch von wesentlicher Bedeutung. Er meint, „dass man eigentlich immer von so einer Art gutmütigem Agenten, sagen wir mal so, ausgeht. Also nicht wie in der menschlichen Gesellschaft, dass manche Leute nicht arbeiten wollen.“ (UW3, 14: 26-29) Ein anderer Gesprächspartner verweist darauf, dass Koordination mittels Hilfsbereitschaft in vielen Situationen in funktionaler Hinsicht von Vorteil ist. Dies könne man sich deutlich machen, wenn man sich beispielsweise vorstelle, die Mitglieder eines Operationsteams im Krankenhaus würden über die wechselseitige Bereitschaft zu Handreichungen jeweils miteinander verhandeln müssen: „Wenn die versuchen, irgendwie zu einem Kompromiss zu kommen, dann ist der Patient schon längst tot. Also muss man sich überlegen, dass man sagt, man geht davon aus, dass

einfach alle Agenten, die immer wieder daran beteiligt sind, von sich aus konstruktiv sind. Oder einfach alles, was sie wissen, bereitstellen.“ (UW7, 8: 26-33)

Es sind solche erwünschten Vereinfachungen von Abstimmungsprozessen, die im Effekt dazu führen, dass Agenten in Multiagentensysteme in vielerlei Hinsichten in einer zumeist nicht weiter explizierten Weise als kooperationsbereit konzipiert werden, also beispielsweise nicht in der Lage sind, Informationen vorzuenthalten, um die sie gebeten werden, Aussagen wider besseres Wissen zu treffen, ihre Problemlösungsfähigkeit strategisch zurückzuhalten oder Ähnliches.

Das Kontraktnetz-Protokoll

Der gegenwärtig in der Verteilten Künstlichen Intelligenz geläufigste Koordinationsmechanismus ist jedoch das Kontraktnetz-Protokoll. Über diesen Koordinationsmechanismus sagt einer unserer Gesprächspartner, „dass das eigentlich nach wie vor das ist, was alle Leute, wenn der Name VKI fällt, zuerst sagen: Das ist es.“ (IW5, 14: 42-45) Das Kontraktnetz-Protokoll ist der idealisierten Interaktionssituation des Markttausches bei nachfrageorientierten Märkten nachgebildet. Es beruht auf drei Aktionen: der Ausschreibung, dem Gebot und dem Zuschlag. Mittels des Kontraktnetz-Protokolls kann ein Agent, der eine Aufgabe delegieren möchte, seine Nachfrage anderen Agenten bekannt geben, indem er eine Ausschreibung vornimmt. Diejenigen Agenten, die an der Übernahme der Aufgabe interessiert sind, senden darauf ein Gebot, in dem sie mitteilen, zu welchen Konditionen sie zur Ausführung bereit und in der Lage sind. Der nachfragende Agent wählt das für ihn geeignetste Gebot aus und gibt dem betreffenden Agenten den Zuschlag, womit dieser die Ausführungsverantwortung übernimmt (vgl. Davis/Smith 1983: 77ff.).

Für ein breites Feld unterschiedlicher Anwendungen werden Multiagentensysteme konstruiert und evaluiert, die das Kontraktnetz-Protokoll als den grundlegenden Koordinationsmechanismus vorsehen. Die prominentesten dieser angezielten Anwendungsfelder sind die Fertigungssteuerung, bei der es dann darum geht, die Bearbeitung von Werkstücken auf diese Weise auf Maschinen zu verteilen, und die Transportlogistik, wo die Zielsetzung darin besteht, Speditionsaufträge so auf die einzelnen Transporteinheiten zu verteilen, dass bei möglichst hoher Auslastung die möglichst kürzeste Wegstrecke zurückgelegt wird. In jüngerer Zeit gewinnt die agentenunterstützte Abwicklung von *electronic commerce* zunehmende Bedeutung, wofür, so ein Gesprächspartner, „Marktplatz logischerweise die beste Metapher“ (UW12, 6: 5) sei.

Die Eignung des Kontraktnetz-Protokolls als technisches Verfahren der Agentenkoordination wird von einem unserer Gesprächspartner besonders prägnant herausgestellt. Er sagt: „*Contract net* ist ein Koordinationsverfahren, wo Sie im Prinzip eine Aufgabe auf andere Agenten verteilen. ... Beim *contract net* ist es im Prinzip so ...: Einer sagt: ‚Ich habe ein Problem.‘ Der schreibt das an bestimmte Leute aus, die geben ihm halt ein Angebot ab. D.h. die sagen, wie sie das Problem lösen, und er wählt dann halt das beste aus. Das ist letztendlich nichts anderes als ein *task distribution*-Verfahren. Irgendwie eine Aufgabe an andere verteilen. Da habe ich Kriterien, nach denen ich das mache, das sind dann halt meine Optimierungskriterien, die wende ich darauf an.“ (IW1, 9: 40 - 10: 4)

Mit Sozialität in einem emphatischen Sinne, dergestalt, dass sich die Beteiligten über ein gemeinsames Vorgehen verständigen müssten, einander wechselseitig überzeugen wollten oder gegebenenfalls einen Kompromiss suchen, mit dem alle Seiten leben können, habe dieses Koordinationsverfahren mithin nichts zu tun. Die Entscheidung, welcher Agent welche Aufgabe abgibt oder übernimmt, erfolgt vielmehr nach berechenbaren Optimalitätskriterien: „Die

Systeme, die wir jetzt haben, ... brauchen ... nicht zu verhandeln, sondern können halt Informationen austauschen. ... Ein einfaches Beispiel ist: Ich will halt eine Maschine finden, für die Verarbeitung. Ich frage zwei Maschinen. Also ich weiß, was gemacht werden soll, die beiden Maschinen wissen, was sie können. Und ich frag jetzt die Maschinen. Und die eine Maschine sagt: ‚Ich mache das in zehn Sekunden.‘ Und die andere sagt: ‚Zwölf.‘ Das Entscheidungskriterium ist: So schnell wie möglich durch die Fertigung durch. Dann ist klar, ich muss die Maschine mit zehn Sekunden nehmen.“ (IW1, 13: 8-24) Als technisches Problemlösungsverfahren, so die Aussageintention dieses Gesprächspartners, habe das Kontraktnetz-Protokoll seine Herkunft aus der Welt des Sozialen vollständig abgeschüttelt: „Sicher auch dieses *contract net* ist sicher auch mal motiviert, wie es halt in Organisationen funktioniert oder wie Ausschreibungsverfahren funktionieren. Aber was rausgekommen ist, ist eine Software-Technik.“ (IW1, 10: 19-22)

Ich komme damit zurück zu der Frage, zu deren Beantwortung ich den kurzen Exkurs über die Gutwilligkeitsannahme und das Kontraktnetz-Protokoll vorgenommen habe: Auf welche Weise wird Wissen über die Welt des Sozialen in technisches Problemlösungswissen übersetzt? Vor allem zwei Vorgehensweisen scheinen hier explizit oder unausgesprochen zum Zuge zu kommen: funktionale Rekonstruktion und Abstrahierung von der potenziellen Vielfältigkeit menschlicher Praktiken, Motivlagen und Überzeugungen. Aus der Perspektive funktionaler Rekonstruktion werden wiederkehrende Muster sozialer Verhaltenskoordination vom Resultat her als effizient mit Blick auf bestimmte Situationen verteilten Problemlösens angesehen. Gleichzeitig wird das Verhalten der beteiligten Akteure in einer hochgradig idealisierten Weise konzeptuell reformuliert: Es wird unterstellt, dass sich die Akteure so verhalten, wie sie sich verhalten müssten, um dem fraglichen Koordinationsmuster in seiner Funktionalität uneingeschränkt zur Durchsetzung zu verhelfen.

Eine solche funktionale Rekonstruktion liegt etwa vor, wenn Gutwilligkeit unter dem Gesichtspunkt der Einsparung an Einigungsaufwand als effizienter Koordinationsmechanismus beschrieben wird (vgl. das obige Zitat UW7, 8: 26-33). Zugleich impliziert dies eine hochgradige Abstraktion von den Befindlichkeiten der beteiligten Individuen, deren Gutwilligkeit nun kontrafaktisch so behandelt wird als wäre sie motiviert durch einen solchen Effizienzgewinn und nicht, wie dies im empirischen Fall sehr viel wahrscheinlicher ist, durch spezifische Solidaritäten oder Wertorientierungen. Der gleiche Sachverhalt funktionaler Rekonstruktion kommt auch in dem folgenden Zitat Hubermans zum Ausdruck. Seine These lautet, dass die Konzepte der Wirtschaftswissenschaften bei der Koordination künstlicher Agenten zukünftig möglicherweise eine größere Rolle spielen werden als in der realen Ökonomie, weil Menschen „viel zu irrational für die arg mechanischen Lehren der Volkswirte“ seien, Agenten täten „dagegen genau das, was man ihnen vorschreibt“ (Siegele 1997; vgl. auch Huberman/Clearwater 1995).

Dieses letzte Zitat legt die Vermutung nahe, dass sich die Multiagentensystem-Forschung für solche sozialwissenschaftlichen Konzepte besonders interessieren müsste, die bereits selbst funktionale Rekonstruktionen sozialer Gegebenheiten darstellen, insbesondere, wenn damit zugleich der Anspruch verbunden wird, auf der gewonnenen konzeptionellen Grundlage entsprechende Koordinationsformen dann auch einrichten zu können. Den noch günstigeren Bezugspunkt müssten darüber hinaus solche Konzepte bieten, die zur gezielten Einrichtung in bestimmten Hinsichten effizienter Koordinationsmechanismen überhaupt erst erdacht worden sind.

In der Tat finden wir in der einschlägigen Literatur etliche Überlegungen, die in diese Richtung gehen. Einige Beispiele sind: (1) Die so genannte „scientific community metaphor“ (vgl. Kornfeld/Hewitt 1981): Ausgehend von der Feststellung, dass wissenschaftliche Gemeinschaften sich „als außerordentlich erfolgreich beim Problemlösen“ (Kornfeld/Hewitt 1981: 311) erwiesen hätten, wird hier der Versuch unternommen, ein vereinfachtes und idealisiertes Modell charakteristischer Merkmale wissenschaftlicher Gemeinschaften zu formulieren und softwaretechnisch umzusetzen. Dabei rekurren Kornfeld und Hewitt auf Aussagen unter anderem von Popper und Lakatos über bestimmte strukturelle Bedingungen des Erfolgs wissenschaftlicher Forschung. (2) Die Analogie zwischen verteilten Systemen und menschlichen Organisationen (vgl. Fox 1981: 70ff.): Hier geht es darum, unter Rückgriff auf organisations-theoretisches Wissen idealtypische Organisationsstrukturen zu unterscheiden, ihre Vor- und Nachteile mit Blick auf unterschiedliche Problemstellungen zu identifizieren und die für die jeweilige Problemstellung angemessenste Organisationsstruktur im Multiagentensystem nachzubilden (vgl. Kirn 1996; Malone 1990). (3) Der Rekurs auf *economic mechanism design*: Aufgegriffen werden hier Konzepte aus einem noch relativ jungen Feld wirtschaftswissenschaftlicher Forschung, die darauf zielen, Interaktionsregeln für ökonomische Transaktionen zu erfinden, die Akteure von Verhaltensweisen abhalten sollen, die die Gesamteffizienz des fraglichen Tauschzusammenhangs reduzieren. Beispielsweise geht es dabei um die Konstruktion von Verfahrensregeln für Auktionen, die verhindern sollen, dass unangemessen hohe Preise entstehen (vgl. Parkes 2000; Varian 1995).

Zur Kennzeichnung des Referenzbereichs, auf den sich die Multiagentensystem-Forschung zur Bearbeitung des Koordinationsproblems bezieht, habe ich den Begriff „die Welt des Sozialen“ verwendet und sie charakterisiert als Bereich des zusammenwirkenden Verhaltens oder Handelns von Lebewesen. Im Folgenden habe ich mich dann auf denjenigen Teilbereich beschränkt, auf den die VKI vorrangig rekurriert: den des handelnden Zusammenwirkens menschlicher Akteure. Angesichts des schon erwähnten Umstandes der Vielfältigkeit und Heterogenität der Konzepte und Vorstellungen über diesen Phänomenbereich sowie angesichts dessen, dass die Bezugnahme – als Konzepttransfer – stets vermittelt durch solche Konzepte und Vorstellungen stattfindet, ist mit dieser Bestimmung des Referenzbereichs noch keine hinreichend präzise Eingrenzung erreicht. Denn eine Antwort auf die Frage, welche und wessen Konzepte in der Multiagentensystem-Forschung zum Tragen kommen, ist damit noch nicht gegeben. Einige diesbezügliche Hinweise finden sich allerdings bereits in den eben präsentierten Befunden.

Wessen Konzepte sind es, auf die VKI-Forscher typischerweise zuerst stoßen, wenn sie auf die Suche nach Koordinationsmechanismen gehen? Ganz überwiegend sind es nicht die Wissensbestände der zuständigen wissenschaftlichen Fachdisziplinen. Weder ist das Kontraktnetz in Auseinandersetzung mit wirtschaftswissenschaftlichen Theoriebeständen entwickelt worden noch liegen der Gutwilligkeitsannahme soziologische Konzepte normorientierten Verhaltens zu Grunde. Es handelt sich zunächst vielmehr um Alltagstheorien, also um Vorstellungen über soziale Abstimmungsprozesse, die sich die VKI-Forscher als Teilnehmer und Beobachter ihres gesellschaftlichen Alltags bilden. Ein Beispiel hierfür ist der Rekurs auf eigene Erfahrungen bei der Organisation von Forschungszusammenhängen, den einer unserer Gesprächspartner benennt: „(Ich hab’ mich schon mein Leben lang mit Wissensorganisation beschäftigt, ich versuche, meinen Bereich hier gut zu organisieren Und alle diese verschiedenen Ideen, wie man eigentlich in einem Konzert mit mehreren Beteiligten zusammenarbeiten kann, kommen jetzt bei unseren Agentensystemen auch vor, in der Organisation der Problemlösung.“ (UP2, 1: 50-2: 4)

Aber auch dort, wo Wissensbestände sozialwissenschaftlicher Fachdisziplinen nun doch ins Spiel kommen, erfolgt die Bezugnahme zunächst wenig systematisch. Sie ergibt sich eher koinzidentuell, wie dies die folgende Äußerung dokumentiert: „(W)ir mussten zum Beispiel im Vordiplom Wirtschaftswissenschaften alle ... die grundlegenden Unternehmensformen kennen. ... Und eigentlich lag für mich, als ich diese ganzen Stücke dann zusammengebracht habe für (das fragliche Multiagentensystem, Anm. d. Verf.), sofort auf der Hand: Das ist wie so'n Projektteam. Ich hab' dann auch da gar nicht mehr groß ... weitergesucht, weil, zunächst mal war natürlich erst das Wichtige: Wir müssen zeigen, dass es auf dem Rechner funktioniert. ... Und als das geklappt hat, naja, hat das Ganze für sich selbst gesprochen.“ (UW1, 18: 1-16)

Der Rekurs auf das eigene Alltagswissen und die eigene Allgemeinbildung bezüglich sozialer Abstimmungsprozesse lieferte in den ersten Jahren einen großen Teil der Konzepte für die Multiagentensystem-Forschung und er spielt auch heute noch eine wichtige Rolle. Inzwischen gewinnt aber auch die Bezugnahme auf sozialwissenschaftliche Wissensbestände an Bedeutung. Dafür scheint es insbesondere zwei Gründe zu geben. Der eine Grund hängt eng mit der Frage zusammen, welche Konzepte sozialer Verhaltensabstimmung für die VKI als einem primär an technologischen Problemlösungen orientierten Unterfangen von Interesse sind. Wie oben gezeigt, sind dies insbesondere solche Konzepte, die das handelnde Zusammenwirken von Akteuren unter dem Gesichtspunkt ihrer Funktionalität für die Erreichung bestimmter übergreifender Ziele analysieren, und besser noch: Konzepte, die auf die Einrichtung darauf gerichteter Mechanismen der Verhaltenskoordination zielen. Entsprechende normative Ansätze der betriebswirtschaftlichen Organisationstheorie etwa oder des *economic mechanism design* sind für die VKI besonders attraktiv und werden zunehmend aufgegriffen, weil sie schon im Referenzbereich als Techniken, nämlich als berechenbare Verfahren der Verhaltensabstimmung konstruiert wurden, was eine Übertragung auf Multiagentensysteme außerordentlich begünstigt. Ein zweiter Grund für eine verstärkte Bezugnahme auf sozialwissenschaftliche Konzepte ergibt sich aus der Umorientierung von Problemähnlichkeit auf Strukturgleichheit, auf die ich im nun folgenden Abschnitt eingehe.

4. Strukturgleichheit: Die Welt des Sozialen als Bezugsquelle viel versprechender Problemstellungen

Konzeptübertragungen führen nicht nur zu einer neuartigen und unter Umständen Gewinn bringenden Betrachtungsweise des interessierenden Forschungsgegenstandes. Sie bringen zugleich eine eigentümliche Veränderung der Wahrnehmung des Referenzbereiches mit sich. Auf diesen Punkt hat bereits Black hingewiesen und ihn am Beispiel der Metapher „Der Mensch ist ein Wolf.“ illustriert: „Wenn die Bezeichnung Wolf einen Menschen in ein bestimmtes Licht rückt, so darf man darüber nicht vergessen, daß die Metapher den Wolf dabei menschlicher als sonst erscheinen läßt.“ (Black 1954/1983: 75) Man betrachtet den Referenzbereich eben unter dem Gesichtspunkt seiner Nützlichkeit für Problemstellungen, die innerhalb eines anderen Bereichs generiert worden sind, und interessiert sich dementsprechend nur sehr selektiv für den dortigen Wissensbestand.¹¹ Erscheint in Blacks Beispiel der Wolf menschlicher als sonst, so erscheint in unserem Fall menschliche Verhaltensabstimmung sehr viel technischer als sonst.

11. Stachowiak (1973: 155f.) spricht in seiner allgemeinen Modelltheorie in diesem Zusammenhang von „präterierten Attributen“ und bezeichnet damit jene Merkmale und Eigenschaften des Referenzkontextes, die im Modell weggelassen werden.

Unter bestimmten Umständen ist eine solche, auf partielle Problemähnlichkeit fokussierte Selektivität offensichtlich hochgradig sinnvoll. Dann nämlich, wenn es darum geht, sich für eine gegebene und präzise formulierte Problemstellung eine potenziell nutzbare Problemlösung aus dem Referenzbereich gleichsam nur abzuholen. Die Konstrukteure des Schreibklaviers beispielsweise mussten sich nicht für jene Aspekte der Klaviermechanik interessieren, die eine Dämpfung des angeschlagenen Tones ermöglichen. Sie konnten auf Grund ihrer Problemstellung von vornherein wissen, dass alles diesbezügliche Konstruktionswissen des Klavierbaus für sie ohne Bedeutung sein würde. In diesen Idealfällen technikbezogener Konzeptübertragung können die Protagonisten eine eventuelle Kritik seitens des Referenzkontextes, die dortigen Konzepte würden unzulässig verkürzt und willkürlich zurechtgebogen, getrost ignorieren. Getrost ignoriert werden kann auch die Forderung, man müsse erst eine eigenständige fachwissenschaftliche Kompetenz in dem Referenzbereich erwerben, bevor man an die Übertragung der dortigen Problemlösungen gehen könne.¹²

Anders sieht die Situation aus, wenn eine nur sehr grob umrissene Problemähnlichkeit die Basis des Konzepttransfers bildet, wie dies bei der Multiagentensystem-Forschung in wesentlichen Hinsichten der Fall ist. Zwar lassen sich auch auf dieser Grundlage nützliche Problemlösungen importieren, wofür die Gutwilligkeitsannahme wie auch das Kontraktnetz-Protokoll Beispiele sind. Aber diese Problemlösungsmechanismen bleiben ähnlich unterspezifiziert wie die Problemstellung selbst.¹³ Das, was das Alltagswissen und die Sozialwissenschaften an potenziell ausbeutbarem Wissen zu bieten haben, ist damit offensichtlich noch lange nicht ausgeschöpft. Dort, wo diese Einschätzung forschungsleitend wird, kommt es zu einer überraschenden Umkehrung des Begründungszusammenhangs der Ähnlichkeitsbeziehung: Statt nach Konzepten Ausschau zu halten, die sich in geeigneter Weise auf gegebene Problemstellungen der Multiagentensystem-Forschung übertragen lassen – ohne sich weiter für die Ähnlichkeit der technischen Lösung mit den korrespondierenden Abläufen in der Welt des Sozialen zu interessieren – wird nun die Herstellung von Ähnlichkeit zwischen sozialweltlichen Phänomenen und Multiagentensystemen selbst zum Designkriterium, ohne dass ein direkter Problembezug der resultierenden technischen Konzepte vorliegt.

Sehr schön lässt sich diese Umorientierung anhand einer Passage eines Textes von Rosen-schein und Genesereth nachvollziehen, der in der Verteilten Künstlichen Intelligenz als Auslöser der Kritik an der Gutwilligkeitsannahme gilt (vgl. Decker 1987: 731). Die Autoren argumentieren: „In der wirklichen Welt sind Handelnde nicht notwendigerweise gutwillig in ihren aufeinander bezogenen Handlungsweisen. Jeder Handelnde hat seinen eigenen Satz von Wünschen und Zielen und wird einem anderen Handelnden nicht notwendigerweise mit Informationen oder Aktionen helfen. Konflikte zwischen Handelnden existieren, natürlich sind sie aber keine totalen Konflikte. Oft gibt es die Möglichkeit für Kompromiss und wechselseitig nutzbringende Aktivität. Die bisherige Forschung in der VKI, soweit sie von der Annahme gutwilliger Agenten ausging, ist im Allgemeinen unfähig gewesen, diese Formen von Interak-

12. Für die Multiagentensystem-Forschung ist die entsprechende Forderung insbesondere von Gasser (1991: 111f.) erhoben worden (vgl. zustimmend Malsch 1997: 16).

13. Einer unserer Gesprächspartner benennt seine diesbezügliche Unzufriedenheit mit dem Kontraktnetz-Protokoll wie folgt: „Der Punkt bei dem Kontraktnetz ist, das sagt dir im Prinzip: Gut, ich habe irgendeinen Job zu tun, ich frage die Leute (d.h. die Agenten, Anm. d. Verf.). Und die Leute geben mir eine Zahl zurück und der mit der höchsten Zahl gewinnt. ... Das eigentliche Problem ist doch: Ja, wie wird denn diese Zahl berechnet? Und da ... kriegt man eigentlich viel zu wenig Hilfe. ... Das Kontraktnetz ist schön und gut, aber es löst dir allein nicht die Probleme. Gibt dir nur einen allgemeinen Mechanismus, wie du eben Aufgaben in einer verteilten Umgebung an die beste Partie (verteilen kannst).“ (IW5, 14: 39 - 15: 3)

tion zu behandeln. Für intelligente Agenten, die in der Lage sind zu interagieren, auch wenn ihre Ziele nicht identisch sind, würde es vielfältige Verwendungsmöglichkeiten geben.“ (Rosenschein/Genesereth 1985/1988: 227)

Der springende Punkt ist, dass nicht eine vermutete oder erwiesene technische Ineffizienz oder Unterlegenheit der Grund für die Kritik an der Gutwilligkeitsannahme ist. Das zentrale Argument ist vielmehr, dass diese das Verhalten von Handelnden in der wirklichen Welt nicht zutreffend abbilde. Ein vergleichbarer Einwand im Fall des Schreibklaviers, der da hätte lauten können, dass das feine Differenzierungsvermögen der Erzeugung unterschiedlich lauter Töne bei der Hebelmechanik der Schreibmaschine außer Acht gelassen werde, hätte mit einiger Sicherheit kein Gehör gefunden. Warum also kann die entsprechende Forderung hier zu einem entscheidenden Argument gegen die Güte des importierten Konzeptes werden? Ein Hinweis zur Beantwortung dieser Frage findet sich am Ende des eben angeführten Zitats: Ausschlaggebend ist die Vermutung, dass eine originalgetreuere Nachbildung sinnvoll wäre, weil sich daraus Lösungen für noch nicht weiter spezifizierte Probleme ergeben könnten. Im Fall des Schreibklaviers kann dieses überschießende Problemlösungspotenzial dagegen umstandslos abgeschnitten werden, weil absehbar ist, dass es für die Bearbeitung des gegebenen Problems nichts beitragen kann.

Um es pointiert zusammenzufassen: Die Veränderung im Prozess technikbezogener Konzeptübertragungen, um deren Beschreibung es hier geht, besteht in einer Umorientierung von Problemähnlichkeit auf Strukturgleichheit. Die auf eine unterstellte Problemähnlichkeit gegründete Suche nach Problemlösungen aus dem Referenzbereich wird durch ein gegenläufiges Verfahren der Konzeptübertragung ergänzt oder sogar teilweise abgelöst. Es zielt auf die Konstruktion von Multiagentensystemen, die so strukturiert sind wie bestimmte menschliche Interaktionszusammenhänge.¹⁴ Handlungsleitend ist dabei die zunächst noch unbestimmte Hoffnung, dass man auf diese Weise zu softwaretechnischen Problemlösungsmechanismen gelangen wird, zu denen sich dann – wie im Bereich menschlicher Verhaltensabstimmung so auch in der Informatik – passende Probleme finden werden. Ich möchte diesen Punkt wieder an zwei Konzepten der Multiagentensystem-Forschung illustrieren: der BDI-Architektur und dem Konzept offener Systeme.

14. Die Annahme, dass sich die interessanteren und nutzbringenderen Lösungen ergeben können, wenn man bei der Konstruktion der Agenten fragt: „Wie machen’s denn die Menschen?“, also auf der Grundlage von Strukturgleichheit analogisiert, kommt in der folgenden Interviewpassage sehr schön zum Ausdruck. Es geht dabei um Agenten, die in verteilten, heterogenen Informationssystemen nach Informationen suchen: „Der eigentliche Zugangspunkt war der: Informationsagenten werden bezahlt, müssen dem anderen was geben. Diese(r) Transaktionskosten(ansatz) ist langweilig. ... Vor allem wenn sie autonom sind, werden die ja wohl nicht alles akzeptieren. Machen wir ja auch nicht: Wenn wir miteinander uns unterhalten oder miteinander handeln würden, auf dem Marktplatz, geht das ja genauso. So und da ist man schon schwupp bei Mechanismen oder Methoden: Wie machen’s denn die Menschen? BWL, Betriebswirtschaft, Entscheidungstheorie – ganz klassisches Gebiet. So, und da guckt man. Und eins davon ist die kooperative Spieltheorie und das hat mich besonders fasziniert, weil da nämlich Koalitionen gebildet werden. Und Koalitionen haben mich ganz stark gefesselt. Weil dieser Begriff Koalition nämlich sowohl nutzbringend sein kann, was dann ja auch wieder den Verkaufswert (Lachen) für Produkte ausmacht, und auf der anderen Seite auch eine Art von: So könnte ich mich eventuell selbst verhalten.“ (UW2, 21: 45 - 22: 22) Interessant ist für unseren Zusammenhang auch die Ablehnung des Transaktionskostenansatzes: Den in den Wirtschaftswissenschaften für entsprechende Problemstellungen etablierten Lösungsansatz einfach zu importieren (Rekurs auf Problemähnlichkeit), empfindet unser Gesprächspartner als „langweilig“.

Die BDI-Architektur

Die Konsequenz, die Rosenschein und Genesereth aus ihrer Kritik der Gutwilligkeitsannahme ziehen, besteht darin, die Konstruktion von Agenten ins Auge zu fassen, die je eigene Ziele besitzen und die über bestimmte Strategien der Zielerreichung verfügen, welche die möglichen Interdependenzen eigener Verhaltensoptionen mit denen anderer Agenten berücksichtigen. Die BDI-Architektur („BDI“ steht für „belief, desire, intention“) ist Anfang der Neunzigerjahre in Weiterführung entsprechender Überlegungen entwickelt worden (vgl. Cohen/Levesque 1990; Rao/Georgeff 1991; Shoham 1993).

Einer unserer Gesprächspartner rekonstruiert die Überlegungen, die zur Entstehung der BDI-Architektur geführt haben, folgendermaßen: „Nun ist also die Frage: Was ist die geeignetste Technologie für Programme, die in irgendeiner Weise autonom sein sollen, die miteinander kommunizieren können, kooperieren sollen und auch vielleicht in der Lage sein sollen zu lernen? ... Wenn ich jetzt sage, das Programm soll also selber aktiv sein. Na, was sind Einheiten, die mental irgendwie aktiv sind? Das sind wir Menschen. Also fangen wir jetzt an zu überlegen: ... Was bestimmt das Handeln des Menschen? Erstmal ein bestimmtes Wissen: *belief*. Dann haben wir bestimmte Aufgaben, Aufträge oder auch Ziele, je nachdem. Das sind die berühmten *desires* oder *goals*. Und dann ist da noch irgendwo, wenn es komplexer wird, wenn man mehrere solcher Ziele gleichzeitig verfolgen will ..., dann muss (der Agent) sehen, dass er nicht in Konflikte kommt, dann muss er irgendwie Entscheidungen treffen: *commitments* oder *intentions*, Absichten. Das ist dann die berühmte BDI-Architektur. ... Ja, das ist einfach ein natürlicher Weg, so etwas zu beschreiben.“ (UP1, 4: 10-32)

Die BDI-Architektur basiert auf der Überlegung, dass es bestimmte typische Formen intentionaler Einstellungen – wie Überzeugungen, Wünsche oder Absichten – sind, mit denen menschliche Akteure ihre eigenen Handlungen und diejenigen anderer Akteure im Alltag unproblematisch zu erklären vermögen und die sie der wechselseitigen Koordination ihrer Handlungen zu Grunde legen: „Wenn wir menschliche Aktivitäten erklären wollen, ist es häufig nützlich Aussagen der folgenden Art zu treffen: Janine nahm ihren Schirm, weil sie überzeugt war, dass es zu regnen anfangen würde. Michael arbeitete hart, weil er einen Dokortitel haben wollte. Diese Aussagen bedienen sich einer Alltagspsychologie, mittels derer menschliches Verhalten durch die Attribution von Einstellungen ... vorhergesagt und erklärt wird. Diese Alltagspsychologie ist gut etabliert: Die meisten Leute, die die obigen Aussagen lesen, würden sagen, dass deren Bedeutung für sie völlig klar ist“ (Wooldridge/Jennings 1995: 4). Diese alltagspsychologischen Kategorien werden nun verwendet, um Agenten mit softwaretechnischen Komponenten auszustatten, die zu einem Verhalten führen, das sich gleichermaßen als Ausdruck intentionaler Eigenschaften interpretieren lässt. Eine bekannte Definition des BDI-Agenten lautet dementsprechend: „Ein Agent ist eine Entität, deren Zustand als bestehend aus mentalen Komponenten wie Überzeugungen, Fähigkeiten, Wahlmöglichkeiten und Verpflichtungen angesehen wird. Diese Komponenten sind in präziser Weise definiert und stehen in ungefährender Korrespondenz zu ihren Entsprechungen im Alltagsverständnis.“ (Shoham 1993: 52)

In diesen Überlegungen kommt die Fokussierung auf Strukturgleichheit statt auf Problemähnlichkeit deutlich zum Ausdruck. Die Motivation für diese Fokussierung, so bringt es ein anderer unserer Gesprächspartner auf den Punkt, bestehe darin, „dass man die Agententechnologie begreift als eine Entwurfsmethode, die aus dem normalen Leben resultiert sozusagen. Man versucht also, noch stärker als in der objektorientierten Modellierung, eine viel stärkere Abbildung der menschlichen Welt auf dem Rechner tatsächlich hinzukriegen.“ (UW4, 8: 50 - 9: 1) Ergänzend ist anzumerken, dass zwar sowohl in den Äußerungen unserer Gesprächspartner

wie auch in der zitierten Literatur auf die hohe alltagsintuitive Plausibilität der Konzepte verwiesen wird, die der BDI-Architektur zu Grunde liegen. Dennoch liegt keine direkte Übertragung von Alltagsvorstellungen der VKI-Forscher selbst vor. Vielmehr findet hier ein Rekurs auf sozialwissenschaftliche Wissensbestände statt, nämlich auf sozialpsychologische und philosophische Forschungen zur Alltagspsychologie.¹⁵

Offene Systeme

Auch das von Carl Hewitt (1986; 1991) in die Verteilte Künstliche Intelligenz eingeführte Konzept offener Systeme zielt darauf Strukturgleichheit herzustellen, darauf nämlich, Muster von Interaktionsbeziehungen nachzubilden, die Hewitt für die Strukturierung von Organisationen als wesentlich ansieht. Er rekurriert dabei auf die Organisationstheorie des symbolischen Interaktionismus (vgl. Strübing 1998). Als offene Systeme bezeichnet Hewitt VKI-Systeme, die aus einer nicht von vornherein festgelegten und unter Umständen sehr großen Zahl von Agenten bestehen, welche laufend mit neuen, unerwarteten und inkonsistenten Informationen konfrontiert sind, wobei die Agenten ihre Ziele, Verhaltensoptionen und Wissensbestände wechselseitig zunächst nicht kennen, sondern voneinander nur erfahren, was sie mittels expliziter Kommunikation austauschen. Es gibt außerdem keine zentrale Entscheidungsinstanz. Alle Entscheidungen werden lokal getroffen und können stets zu einem unvorhergesehenen Verhalten des Systems führen (vgl. Hewitt 1986: 320ff.; 1991: 80ff.).

Hewitt selbst gibt nur wenige Hinweise darauf, für welche konkreten gegenwärtigen oder zukünftigen Probleme der Multiagentensystem-Forschung es nützlich sein könne, Multiagentensysteme als offene Systeme zu konzipieren. Erst das Internet und die Entdeckung viel versprechender Anwendungsmöglichkeiten internetbasierter Multiagentensysteme hat dafür gesorgt, dass Konzepte offener Multiagentensysteme zu Lösungsansätzen konkretisierbarer Problemstellungen der Verteilten Künstlichen Intelligenz geworden sind. Die neue Problemwahrnehmung ist dabei bei einzelnen unserer Gesprächspartner so dominant, dass offene Systeme nun geradezu als „die einzige Anwendung von diesen Multiagentensystemen“ (IW5, 10: 13f.) deklariert werden: „In dem *world wide web* hast du eigentlich eine extrem gute Ausgangssituation für die Idee, die hinter (offenen) Multiagentensystemen steht. Weil wenn du heute eine Applikation für das *web* entwickelst, dann musst du davon ausgehen, dass das *web* nicht eine Person unter Kontrolle hat. Du bietest Dienste an, du weißt nicht, wer diese Dienste eigentlich in Anspruch nehmen wird. Und das *web* ist ständig in Bewegung. ... Du hast auch nicht alle Knoten unter Kontrolle. Du kannst im Prinzip Agenten designen, die auf dem *web* etwas anbieten, aber du weißt nicht, wie die Nachfrage sein wird. Und schon gar nicht, wie sich das *web* in zwei Jahren entwickelt hat.“ (IW4, 25: 19- 31)

Der Begriff des offenen Systems wird nun zumeist in einer wesentlich problemorientierteren Weise definiert, nämlich als ein Multiagentensystem, „das du selber nicht komplett in der Hand hast ... Wo du selber keine Standards vorschreiben kannst.“ (IW4, 18: 5-8) Dieses Problem entsteht, weil die interessantesten Internet-Anwendungen der Agententechnologie darauf beruhen, dass eine Vielzahl unterschiedlicher Entwickler und Benutzer je ihre eigenen Agenten losschicken, um Dienste anzubieten oder nachzufragen, sodass es nicht mehr den einen Entwickler oder das eine Entwickler-Team gibt, der bzw. das dafür sorgen kann, dass sich die beteiligten Agenten in aufeinander abgestimmter Weise zueinander verhalten.

15. Insbesondere auf die entsprechenden Überlegungen von Dennett (1987) wird in der Literatur zur BDI-Architektur häufig verwiesen (vgl. z.B. Haddadi/Sundermeyer 1996: 169).

Vor dem Hintergrund dieser beiden, auf Strukturgleichheit abstellenden Konzeptübertragungen kann nun erneut die Frage aufgeworfen werden, welche und wessen Konzepte über die Welt des Sozialen es sind, die zur Diskussion stehen. Es sind offensichtlich eher deskriptive als normative Konzepte, also Konzepte, die zunächst darauf zielen, empirisch vorfindliche Strukturen und Prozesse des handelnden Zusammenwirkens zu beschreiben, und die nicht schon gleich – rekonstruktiv oder konstruktiv – nach deren Funktionalität für Koordinationsprobleme fragen. Es handelt sich also, um auf eine oben verwendete Unterscheidung zurückzukommen, eher um empirisch-wissenschaftliche als um technologische Konzepte. Diese Differenz markiert deutlich den Kern dessen, worum es bei der Umorientierung von Problemähnlichkeit auf Strukturgleichheit geht: Um eine Bezugnahme auf Beschreibungen von Phänomenen des Referenzbereichs, die dort von beträchtlicher Bedeutung zu sein scheinen, ohne dass zugleich ein direkter funktionaler Zusammenhang genannt werden kann, der begründet, weshalb Gleiches auch für das primäre Forschungsfeld gilt.

In gewisser Weise erfolgt mit der Umorientierung auf Strukturgleichheit eine Verselbstständigung der Bezugnahme auf den Referenzbereich, insofern nämlich als die Konzeptübertragungen nun nicht mehr vollständig durch die Problemstellungen des primären Forschungsfeldes kontrolliert werden. In diesen Kontext gehört die oben angesprochene Forderung nach einer sozialwissenschaftlichen Fundierung der VKI, die genau darauf zielt, den Referenzbereich als eine Inspirationsquelle für die Lösung nicht oder unzulänglich spezifizierter Problemstellungen des primären Forschungsfeldes zu erschließen. Maßgeblich ist dabei die Vermutung, dass Phänomene, die innerhalb des Referenzbereichs eine Rolle spielen, auch im primären Forschungsfeld von Bedeutung sind. Diese Vermutung – das jedenfalls ist die Quintessenz der beiden Fallbeispiele – ist die wesentliche Motivation für die Umorientierung auf Strukturgleichheit.¹⁶ Darin, dass sich diese Vermutung als falsch erweisen kann, besteht das wesentliche Innovationsrisiko technikbezogener Konzeptübertragungen, die auf Strukturgleichheit setzen. Umgekehrt besteht die wesentliche Innovationschance darin, auf unerwartete Problemlösungen für den primären Forschungsbereich zu stoßen, das heißt auf solche, die außerhalb des durch die dortigen Problemstellungen definierten Suchhorizontes liegen.

Die Umorientierung auf Strukturgleichheit, um es zusammenfassend festzuhalten, stellt eine Reaktion auf Probleme dar, die sich für die auf Problemähnlichkeit rekurrierende technikbezogene Konzeptübertragung stellen können. Es handelt sich dabei um das Problem, dass die Ähnlichkeitsbeziehung konstituierende Problem unter Umständen zu unspezifisch ist, um die Suche nach sinnvoll übertragbaren Konzepten wirklich instruieren zu können, und/oder umgekehrt um das Problem, dass der durch die Problemstellungen der primären Forschungsfeldes konstituierte Suchhorizont zu eng ist, um potenziell viel versprechende Problemlösungen des Referenzbereichs in den Blick zu bekommen.

Diese Probleme stellen sich im Fall der Multiagentensystem-Forschung aber durchaus nicht für alle Forscher und auch nicht für alle Forscher in gleicher Weise. Es wäre deshalb unzutreffend, eine komplette Umstellung von Problemähnlichkeit auf Strukturgleichheit zu vermuten oder hierfür gar einen bestimmten Zeitpunkt identifizieren zu wollen. Die stärksten Vorbehalte gegen Konzeptübertragungen auf der Grundlage von Strukturgleichheit äußern, wenn auch nicht durchgängig, diejenigen unserer Gesprächspartner, die besonders anwendungsnah arbeiten. Sie argumentieren, dass man zunächst die durch die technischen Problemstellungen

16. Sie ist auch eine entscheidende Triebkraft für die interdisziplinäre Kooperationen zwischen Soziologie und Multiagentensystem-Forschung, wie sie durch den DFG-Forschungsschwerpunkt Sozionik (vgl. Malsch et al. 1996; Malsch 2001) und dessen u.s.amerikanische Vorläufer (Bendifallah et al. 1988) repräsentiert werden.

definierten Forschungsaufgaben bewältigen müsse: „Das Software-Engineering für Agentensysteme steckt immer noch in den Kinderschuhen. ... Die meisten Leute interessieren sich ja für irgendwelche komplexen Koordinations- oder Verhandlungsverfahren ... Bevor man jedoch die untere Ebene nicht in den Griff bekommt, dann brauchen wir uns mit den komplexeren Dingen nicht auseinandersetzen.“ (IW1, 4: 10-20) Umgekehrt kann daraus aber nicht der Schluss gezogen werden, dass die Umorientierung auf Strukturgleichheit zugleich auch eine Abkehr von technologischen und eine Hinwendung zu rein wissenschaftlichen Problemstellungen impliziere. Zwar ist richtig, dass Strukturgleichheitsannahmen am deutlichsten von universitären Forschern verwendet werden. Aber auch ihre Forschungsanstrengungen sind, wie oben bereits dargestellt, ganz überwiegend – wenn auch indirekter und mit längerem Atem – auf technologische Problemlösungen hin orientiert. Dies berechtigt zu der These, dass es letztlich auch hier, nur eben über den Umweg der Strukturgleichheit, um die Herstellung von Problemähnlichkeit geht.

5. Ergebnis: Problemähnlichkeit auch über den Umweg von Strukturgleichheit

Ausgangspunkt der voranstehenden Überlegungen war die Annahme eines idealtypischen Falls technikbezogenen Wissenstransfers. Die zentrale These war, dass Strukturgleichheit als Grundlage der Formulierung von Korrelationsregeln zwischen dem primären Forschungsfeld und dem Referenzbereich im Fall technikbezogener Konzeptübertragungen abgelöst wird durch Problemähnlichkeit. Am Beispiel des Schreibklaviers habe ich versucht plausibel zu machen, dass sich einfache Fälle technikbezogener Konzeptübertragungen nach diesem Muster in der Tat gut erklären lassen. Das Beispiel der Multiagentensystem-Forschung hat gezeigt, dass dies auch für bestimmte Aspekte weniger einfach strukturierter Fälle gilt. Dann aber wurde deutlich, dass der Rekurs auf Ähnlichkeit zu gegebenen Problemen des primären Forschungsfeldes hier nicht hinreichend ist und dass in dieser Situation das maßgebliche zusätzliche Kriterium wiederum Strukturgleichheit ist. Folglich stellt sich die Frage, ob das anfangs entwickelte Konzept technikbezogenen Wissenstransfers nur in besonders günstigen Fällen trägt oder ob es sich in sinnvoller Weise so erweitern lässt, dass auch der zusätzliche Gesichtspunkt des Umwegs über Strukturgleichheit Berücksichtigung findet.

Im Vorangegangenen habe ich bereits für die zweite Option votiert, sodass es jetzt abschließend darum gehen muss, diese Entscheidung noch einmal ausdrücklich zu begründen. Hierfür ist es zunächst wichtig sich klar zu machen, dass Strukturgleichheit im Kontext technikbezogenen Konzepttransfers etwas anderes ist als die von Black für wissenschaftliche Konzeptübertragungen als maßgeblich betrachtete Strukturgleichheit. Annahmen über Strukturgleichheit, wie sie den theoretischen Modellen wissenschaftlicher Konzepttransfers zu Grunde liegen, sind einer erkenntnistheoretisch wie immer zu qualifizierenden Bewährung an empirisch beobachteten Phänomenen ausgesetzt. Sie können sich also als erklärungskräftiger oder weniger erklärungskräftig erweisen. Die technische Nachbildung bestimmter Strukturen eines Referenzbereichs hängt in dieser Hinsicht dagegen völlig in der Luft. Ob es sinnvoll ist, eine bestimmte Strukturgleichheit herzustellen, kann nicht anhand gegebener Phänomene des primären Forschungsfeldes überprüft werden, weil die fraglichen Gegebenheiten – technologische Regeln und technische Artefakte – ja erst auf der Grundlage der betreffenden Strukturgleichheitsunterstellung konstruiert werden. Strukturgleichheit ist anders als im Fall wissenschaftlicher Konzeptübertragungen mithin nie eine ausreichende Grundlage für die Herstellung einer Ähnlichkeitsbeziehung, bzw. sie ist es nur dann, wenn man die erfolgreiche technische Nachbildung von Strukturmustern bestimmter Referenzbereiche selbst zum Bewer-

tungskriterium macht. Damit ist dann aber noch nichts darüber ausgesagt, wozu solche Nachbildungen gut sind.

Sobald man davon ausgeht, dass technische Konstruktionen ihre spezifische Bedeutung dadurch erlangen, dass sie als Lösungen für Probleme betrachtet werden, kommt man aus diesem Grund unweigerlich auf Problemähnlichkeit als Basis technikbezogener Konzeptübertragungen zurück. Die Annahme einer Strukturgleichheit zwischen bestimmten Phänomenen eines Referenzbereichs und den erst noch zu konstruierenden Technologien und Techniken des primären Forschungsfeldes erweist sich unter diesen Bedingungen als eine indirekte Form der Herstellung von Problemähnlichkeit. Der Umweg über Strukturgleichheit dient dabei dazu, eine zunächst noch unscharf formulierte Problemähnlichkeit zu spezifizieren. Auch in Fällen, die sich nicht umstandslos in das idealtypische Muster technikbezogener Konzeptübertragungen einfügen lassen, ist also die wahrgenommene Problemähnlichkeit die konstitutive Grundlage der Analogierelation. Zusätzlich aber wird sie im günstigen Fall über den Umweg der Unterstellung von Strukturgleichheit präzisiert, modifiziert oder erweitert.

Mit zwei Schlussbemerkungen möchte ich auf den innovationstheoretischen Ausgangspunkt der hier präsentierten Überlegungen zurückkommen: Aus der eingangs angesprochenen sozialkonstruktivistischen Perspektive sind wissenschaftliche oder technische Neuerungen das Ergebnis sozialer Aushandlungsprozesse, innerhalb derer die involvierten Akteure ihre jeweiligen Ressourcen (z.B. rhetorische Fähigkeiten) einsetzen, um der eigenen Position zur Durchsetzung zu verhelfen. Auch für das Innovationsgeschehen der Multiagentensystem-Forschung lässt sich eine entsprechende Abhängigkeit der Themenkonjunkturen von den Einflussmöglichkeiten ihrer Proponenten beobachten. Sehr illustrativ ist in diesem Zusammenhang, was einer unserer Gesprächspartner über Carl Hewitt und dessen Konzept offener Systeme sagt: „Also der Carl Hewitt, das ist einfach ein Visionär. ... Der Carl Hewitt hat ja auch immer tolle Begriffe erfunden, wie *open systems* ... Ja, und hat das dann eben mit seiner penetranten Stimme da vorgetragen und denn haben wir gesagt: Super! Nicht, es gibt ja so Guru-Typen. Und das ist eben, der Carl Hewitt ist eben so ein Guru-Typ. Und das war auch völlig okay. Weil: Der hat einfach die Personen zum Denken angeregt“ (IW3, 11: 5-11). Zwischen der Annahme, dass es eines „Guru-Typen“ bedurfte, damit das Konzept offener Systeme in der VKI-Forschung die festzustellende Bedeutung erlangen konnte, und der obigen Darstellung dieses Konzepts als Resultat einer Konzeptübertragung besteht dennoch kein Widerspruch. Denn man möchte ja auch gerne wissen, wie Hewitt auf das Konzept offener Systeme gekommen ist, und stößt dann unweigerlich auf dessen Bezugnahme auf den symbolischen Interaktionismus.

Legt man zum anderen die eingangs ebenfalls angesprochene Unterscheidung zwischen inkrementellen und radikalen Innovationen zu Grunde und nimmt man die Binnenperspektive der Informatik ein, so stellen sich Multiagentensysteme sehr viel eher als radikale, denn als inkrementelle Innovation dar: als wesentliche Neuerung, die darin besteht, intelligentes Problemlösen als Resultat des aufeinander abgestimmten Verhaltens einer Mehrzahl von Problemlösern anzusehen, wodurch das zentrale Forschungsparadigma der klassischen KI – die Entwicklung intelligenter Techniken nach dem Vorbild des menschlichen Gehirns – zumindest relativiert wird. In diesem Sinne nehmen die beteiligten Forscher die Multiagententechnologie durchaus als etwas Neues wahr und das ist für die wissenschaftliche Attraktivität dieser Forschungsrichtung nicht unwichtig. Der Begriff des Agenten, so formuliert es einer unserer Gesprächspartner prägnant, kann geradezu als Kurzform der Aussage verstanden werden: „Ich bin jetzt an einer Ecke angelangt, wo ich mit der klassischen KI ... an der Stelle nicht mehr weiterkomme“ (UW7, 39: 19f.). Im Rahmen der KI-Forschung sind Multiagentensys-

teme also mehr als eine bloße Weiterentwicklung und Verbesserung vorhandener Technologien. Sie stehen vielmehr für eine diskontinuierliche Entwicklung im Innovationsgeschehen der Informatik. Es lässt sich aber auch noch eine andere Perspektive einnehmen, eine Perspektive, die auf die Gemeinsamkeiten zwischen der Multiagentensystem-Forschung und der Welt des Sozialen als dem gewählten Referenzbereich abstellt. Und aus dieser Perspektive, die sich für das zusammenwirkende Verhalten selbstständig verhaltensfähiger Einheiten interessiert, wird man überwiegend auf Bekanntes stoßen, wenn man Multiagentensysteme betrachtet. Aus diesem Grund kann, wie in dem vorliegenden Beitrag geschehen, eine kontinuierkeitsbezogene Erklärung in Anschlag gebracht werden, auch wenn sich das Innovationsgeschehen aus der Binnenperspektive – zu Recht – als diskontinuierliche Entwicklung darstellt.

Literatur

- Bendifallah, S./Blanchard, F./Cambrosio, A./Fujimura, J./Gasser, L./Gerson, E.M./Henderson, A./Hewitt, C./Scacchi, W./Star, S.L./Suchman, L./Trigg, R., 1988: The Unnamable: A White Paper on Socio-Computational 'Systems', Research Note 60, DAI Group, University of Southern California, unpublished draft manuscript available from Les Gasser, Department of Computer Science, University of Southern California, Los Angeles.
- Bijker, W.E., 1995: Of Bicycles, Bakelite, and Bulbs: Toward a Theory of Sociotechnical Change. Cambridge, Mass. u.a.: MIT Press.
- Black, M., 1954/1983: Die Metapher. S. 55-79 in: A. Haverkamp (Hrsg.), Theorie der Metapher. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Black, M., 1962: Models and Archetypes. S. 219-243 in: M. Black (Hrsg.), Models and Metaphors. Studies in Language and Philosophy. Ithaca u.a.: Cornell University Press.
- Black, M., 1977/1983: Mehr über die Metapher. S. 379-413 in: A. Haverkamp (Hrsg.), Theorie der Metapher. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Bond, A.H./Gasser, L. (Hrsg.), 1988: Readings in Distributed Artificial Intelligence. San Mateo, Ca.: Morgan Kaufmann.
- Cammarata, S./McArthur, D./Steeb, R., 1983/1988: Strategies of Cooperation in Distributed Problem Solving. S. 102-105 in: A.H. Bond/L. Gasser (Hrsg.), Readings in Distributed Artificial Intelligence. San Mateo, Ca.: Morgan Kaufmann.
- Cohen, P.R./Levesque, H.J., 1990: Intention Is Choice with Commitment. Artificial Intelligence 42: 213-261.
- Collins, H.M., 1983: An Empirical Relativist Programme in the Sociology of Scientific Knowledge. S. 85-113 in: K. Knorr-Cetina/M. Mulkay (Hrsg.), Science Observed. Perspectives on the Social Study of Science. London u.a.: Sage.
- Connah, D./Wavish, P., 1990: An Experiment in Cooperation. S. 197-212 in: Y. Demazeau/J.-P. Müller (Hrsg.), Dezentralized AI, Proceedings of the First European Workshop on Modelling Autonomous Agents in a Multi-Agent World. Amsterdam u.a.: Elsevier.
- Danneberg, L./Graeser, A./Petrus, K., 1995: Metapher und Innovation. S. 9-21 in: L. Danneberg/A. Graeser/K. Petrus (Hrsg.), Metapher und Innovation. Die Rolle der Metapher im Wandel von Sprache und Wissenschaft. Bern u.a.: Paul Haupt.
- Davis, R., 1980: Report on the Workshop on Distributed AI. Sigart Newsletter 73: 42-52.
- Davis, R./Smith, R.G., 1983: Negotiation as a Metaphor for Distributed Problem Solving. Artificial Intelligence 20: 63-109.

- Decker, K.S., 1987: Distributed Problem-Solving Techniques: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics* 17: 729-740.
- Dennett, D.C., 1987: *The Intentional Stance*. Cambridge Mass: MIT Press.
- Dosi, G., 1982: Technological Paradigms and Technological Trajectories. A Suggested Interpretation of the Determinants and Directions of Technical Change. *Research Policy* 11: 147-162.
- Durfee, E.H./Lesser, V.R./Corkhill, D.D., 1987: Cooperation through Communication in a Distributed Problem Solving Network. S. 29-54 in: M.N. Huhns (Hrsg.), *Distributed Artificial Intelligence*. Pitman, Ca. u.a.: Morgan Kaufmann.
- Durfee, E.H./Lesser, V.R./Corkhill, D.D., 1989: Trends in Cooperative Distributed Problem Solving. *IEEE Transactions on Knowledge and Data Engineering* I: 63-83.
- Fox, M.S., 1981: An Organizational View of Distributed Systems. *IEEE Transactions on Systems, Man, and Cybernetics* 11: 70-80.
- Freeman, C./Perez, C., 1988: Structural Crisis of Adjustment, Business Cycles and Investment Behaviour. S. 38-66 in: G. Dosi/C. Freeman/R. Nelson/G. Silverberg/L. Soete (Hrsg.), *Technical Change and Economic Theory*. London u.a.: Pinter.
- Gasser, L., 1991: Social Conceptions of Knowledge and Action: DAI Foundations and Open Systems Semantics. *Artificial Intelligence* 47: 107-138.
- Gilfillan, S.C., 1935/1970: *The Sociology of Invention. An Essay in the Social Causes, Ways and Effects of Technic Invention, Especially, as Demonstrated Historically in the Author's 'Inventing the Ship'*. Cambridge Mass. u.a.: MIT Press.
- Haddadi, A./Sundermeyer, K., 1996: Belief-Desire-Intention Agent Architectures. S. 169-210 in: G.M.P. O'Hare/N.R. Jennings (Hrsg.), *Foundations of Distributed Artificial Intelligence*. New York u.a.: John Wiley.
- Hewitt, C.E., 1986: Offices are Open Systems. *ACM Transactions on Office Information Systems* 4: 271-287.
- Hewitt, C.E., 1991: Open Information Systems Semantics for Distributed Artificial Intelligence. *Artificial Intelligence* 47: 79-106.
- Huberman, B.E./Clearwater, S.H., 1995: A Multi-Agent System for Controlling Building Environments. S. 171-175 in: V. Lesser (Hrsg.), *ICMAS-95. Proceedings of the First International Conference on Multi-Agent Systems, June 12 -14, 1995 in San Francisco, California*. Menlo Park, Ca. u. a.: MIT Press.
- Jennings, N.R., 1996: Coordination Techniques for Distributed Artificial Intelligence. S. 187-210 in: G.M.P. O'Hare/N.R. Jennings (Hrsg.), *Foundations of Distributed Artificial Intelligence*. New York u.a.: John Wiley.
- Kirn, S., 1996: Organizational Intelligence and Distributed Artificial Intelligence. S. 505-526 in: G.M.P. O'Hare/N.R. Jennings (Hrsg.), *Foundations of Distributed Artificial Intelligence*. New York u.a.: John Wiley.
- Knie, A., 1989: *Das Konservative des technischen Fortschritts. Zur Bedeutung von Konstruktionstraditionen, Forschungs- und Konstruktionsstilen in der Technikgenese*, WZB, FS II, 89001. Berlin: Wissenschaftszentrum Berlin.
- Knorr-Cetina, K., 1984: *Die Fabrikation von Erkenntnis. Zur Anthropologie der Wissenschaft*. Frankfurt/Main: Suhrkamp.

- Knorr-Cetina, K., 1995: Metaphors in the Scientific Laboratory: Why are they there and what do they do? S. 329-349 in: Z. Radman (Hrsg.), From a Metaphorical Point of View. A Multidisciplinary Approach to the Cognitive Content of Metaphor. Berlin u.a.: de Gruyter.
- Kornfeld, W.A./Hewitt, C.E., 1981: The Scientific Community Metaphor. IEEE Transactions on Systems, Man and Cybernetics 11: 24-33.
- Latour, B., 1987: Science in Action. How to Follow Scientists and Engineers through Society, Cambridge, Mass.: Harvard University Press.
- Lesser, V.R./Corkhill, D.D., 1983: The Distributed Vehicle Monitoring Testbed: A Tool for Investigating Distributed Problem Solving Networks. The AI Magazine 4: 15-33.
- Malone, T.W., 1990: Organizing Information Processing Systems: Parallels between Human Organizations and Computer Systems. S. 56-83 in: S.P. Robertson/W.W. Zachary/J.B. Black (Hrsg.), Cognition, Computation and Cooperation. Norwood, NJ: Ablex.
- Malsch, T., 1997: Die Provokation der "Artificial Societies". Warum die Soziologie sich mit den Sozialmetaphern der Verteilten Künstlichen Intelligenz beschäftigen sollte. Zeitschrift für Soziologie 26: 3-22.
- Malsch, T., 2001: Naming the Unnamable: Socionics or the Sociological Turn of/to Distributed Artificial Intelligence. Autonomous Agents and Multi-Agent Systems 4: 155-186.
- Malsch, T./Florian, M./Jonas, M./Schulz-Schaeffer, I., 1996: Sozionik: Expeditionen ins Grenzgebiet zwischen Soziologie und Künstlicher Intelligenz. Künstliche Intelligenz 2/1996: 6-12.
- Mambrey, P./Paetau, M./Tepper, A., 1995: Technikentwicklung durch Leitbilder. Neue Steuerungs- und Bewertungsinstrumente. Frankfurt/Main u.a.: Campus.
- Martial, F.v., 1992a: Coordinating Plans of Autonomous Agents. Berlin u.a.: Springer.
- Martial, F.v., 1992b: Einführung in die Verteilte Künstliche Intelligenz. Künstliche Intelligenz 1/1992: 6-11.
- Miller, A.I., 1995: Imagery and Metaphor: The Cognitive Science Connection. S. 199-224 in: Z. Radman (Hrsg.), From a Metaphorical Point of View. A Multidisciplinary Approach to the Cognitive Content of Metaphor. Berlin u.a.: de Gruyter.
- Moulin, B./Chaib-draa, B., 1996: An Overview of Distributed Artificial Intelligence. S. 3-55 in: G.M. O'Hare/N.R. Jennings (Hrsg.), Foundations of Distributed Artificial Intelligence. New York u.a.: John Wiley.
- Nelson, R.R., 1988: Institutions Supporting Technical Change in the United States. S. 312-329 in: G. Dosi/C. Freeman/R. Nelson/G. Silverberg/L. Soete (Hrsg.), Technical Change and Economic Theory. London u.a.: Pinter.
- Parkes, D.C., 2000: Optimal Auction Design for Agents with Hard Valuation Problems. S. in: A. Moukas (Hrsg.), Agent Mediated Electronic Commerce: Towards Next Generation Agent Based Electronic Commerce Systems. Berlin u.a.: Springer.
- Pinch, T.J./Bijker, W.E., 1984: The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. Social Studies of Science 14: 399-441.
- Radman, Z. (Hrsg.), 1995: From a Metaphorical Point of View. A Multidisciplinary Approach to the Cognitive Content of Metaphor. Berlin u.a.: de Gruyter.
- Rammert, W., 1998: Giddens und die Gesellschaft der Heizelmännchen. Zur Soziologie technischer Agenten und Systeme Verteilter Künstlicher Intelligenz. S. 91-128 in: T.

- Malsch (Hrsg.), Sozionik. Soziologische Ansichten über künstliche Sozialität. Berlin: Sigma.
- Rao, A.S./Georgeff, M.P., 1991: Modelling Rational Agents within a BDI Architecture. S. in: J. Allen/R. Fikes/E. Sandewall (Hrsg.), Principles of Knowledge Representation and Reasoning : KR ; International Conference. Proceedings ; 22 - 25 Apr 1991 Cambridge, Mass. San Mateo, Ca: Morgan Kaufmann.
- Rosenschein, J.R./Genesereth, M.R., 1985/1988: Deals Among Rational Agents. S. 227-234 in: A.H. Bond/L. Gasser (Hrsg.), Readings in Distributed Artificial Intelligence. San Mateo, Ca.: Morgan Kaufmann.
- Schlese, M., 1995: Software als "Medium der Kommunikation". Zur Rolle von Leitvorstellungen bei der Konstruktion eines wissensbasierten Systems. S. 359-392 in: W. Rammert (Hrsg.), Soziologie und künstliche Intelligenz. Produkte und Probleme einer Hochtechnologie. Frankfurt/Main u.a.: Campus.
- Schulz-Schaeffer, I., 1999: Technik und die Dualität von Ressourcen und Routinen. Zeitschrift für Soziologie 28: 409-428.
- Schulz-Schaeffer, I., 2000a: Sozialtheorie der Technik. Frankfurt/Main u.a.: Campus.
- Schulz-Schaeffer, I., 2000b: Vergesellschaftung und Vergemeinschaftung künstlicher Agenten. Sozialvorstellungen in der Multiagenten-Forschung, Research Reports 3/2000, Hamburg: TU Hamburg-Harburg,
- Shoham, Y., 1993: Agent-oriented Programming. Artificial Intelligence 60: 51-92.
- Siegele, L., 1997: Ein amerikanischer Forscher bringt Computersystemen die Grundregeln der Marktwirtschaft bei - mit verblüffendem Erfolg. in Die Zeit vom 27.6.1997.
- Stachowiak, H., 1973: Allgemeine Modelltheorie. Wien u.a.: Springer.
- Steeb, R./Cammarata, S./Hayes-Roth, F.A./Thorndyke, P.W./Wesson, R.B., 1981/1988: Distributed Intelligence for Air Flight Control. S. 90-101 in: A.H. Bond/L. Gasser (Hrsg.), Readings in Distributed Artificial Intelligence. San Mateo, Ca.: Morgan Kaufmann.
- Strauss, A.L., 1987: Qualitative Analysis for Social Scientists. Cambridge, Mass. u.a.: Cambridge University Press.
- Strübing, J., 1998: Multiagenten-Systeme als "Going Concern". Zur Zusammenarbeit von Informatik und Interaktionismus auf dem Gebiet der Verteilten Künstlichen Intelligenz. S. 59-89 in: T. Malsch (Hrsg.), Sozionik. Soziologische Ansichten über künstliche Sozialität. Berlin: Sigma.
- Van de Velde, W./Perram, J.P., 1996: Preface. S. V-X in: W. Van de Velde/J.P. Perram (Hrsg.), Agents Breaking Away. Proceedings of the 7th European Workshop on Modelling Autonomous Agents in a Multi-Agent World. Berlin u.a.: Springer.
- Varian, H., 1995: Economic Mechanism Design for Computerized Agents. Paper presented at the USENIX Workshop on Electronic Commerce, July 11-12, 1995, New York, NY. <http://www.sims.berkeley.edu/~hal/people/hal/papers.html>.
- Wooldridge, M.J./Jennings, N.R., 1995: Agent Theories, Architectures, and Languages: A Survey. S. 1-39 in: M.J. Wooldridge/N.R. Jennings (Hrsg.), Intelligent Agents. Proceedings of the ECAI-94 Workshop on Agent Theories, Architectures, and Languages. Berlin u.a.: Springer.